

132013

STIC-Biotech/ChemLib

From: Swope, Sheridan
Sent: Wednesday, September 08, 2004 6:14 PM
To: STIC-Biotech/ChemLib
Subject: 09/992,095

For 09/992,095, pls search and interference search:

SID 54 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

PLS scan
this page

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

From: Chan, Christina
Sent: Tuesday, May 10, 2005 4:57 PM
To: Swope, Sheridan; STIC-Biotech/ChemLib
Subject: RE: 09/992,095

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Swope, Sheridan
Sent: Tuesday, May 10, 2005 4:55 PM
To: Chan, Christina
Subject: 09/992,095

Chris, May I have this rushed?

For 09/992,095 please interference search against the .rnpm and .rapm data bases only:

SID 54, residues 1-270, regular search

SID 54, residues 1-270, oligo search

collect the top 300 hits.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

Scan this pg

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: 5/13 _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 14:39:24 ; Search time 24 Seconds
(without alignments)
449.105 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113

Sequence: 1 MHFCGGTLISPEWVLTAAHC.....GVYRVSRFVTWIEGVNRNN 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1113	100.0	810	1	PLMN_HUMAN
2	1047	94.1	810	1	PLMN_MACMU
3	1018	91.5	333	1	PLMN_CANFA
4	950	85.4	1420	1	APOA_MACMU
5	946	85.0	790	1	PLMN_PIG
6	933.5	83.9	4548	1	APOA_HUMAN
7	927	83.3	812	1	PLMN_MOUSE
8	923	82.9	343	1	PLMN_SHEEP
9	905	81.3	338	1	PLMN_HORSE
10	891	80.1	810	1	PLMN_ERIEU
11	878	78.9	812	1	PLMN_BOVIN
12	448	40.3	454	1	TMS3_HUMAN
13	444.5	39.9	453	1	TMS3_MOUSE
14	444	39.9	417	1	HEPS_HUMAN
15	434	39.0	290	1	PR27_HUMAN
16	433	38.9	436	1	HEPS_MOUSE
17	424.5	38.1	324	1	TEST_MOUSE
18	423	38.0	416	1	HEPS_RAT
19	423	38.0	1035	1	ENTK_BOVIN
20	421	37.8	321	1	TRYG_HUMAN
21	419	37.6	273	1	MCT7_MOUSE
22	418.5	37.6	435	1	TMS4_MOUSE
23	418	37.6	1019	1	ENTK_HUMAN
24	417.5	37.5	1069	1	ENTK_MOUSE
25	416	37.4	1034	1	ENTK_PIG
26	414	37.2	437	1	ACRO_RAT
27	414	37.2	811	1	TMS6_MOUSE
28	413	37.1	273	1	TRYT_SHEEP
29	409.5	36.8	231	1	TRYP_PIG
30	408	36.7	311	1	TRYG_MOUSE
31	406.5	36.5	457	1	TMS5_HUMAN
32	406	36.5	273	1	MCT7_RAT
33	406	36.5	415	1	ACRO_PIG

RESULT 1

ID	PLMN_HUMAN	STANDARD;	PRT;	810 AA.
AC	P00747;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].			
GN	PLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND VARIANT ASN-472.			
RX	MEDLINE=90202879; PubMed=2318848;			
RA	Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;			
RT	"Characterization of the gene for human plasminogen, a key proenzyme in the fibrinolytic system.";			
RL	J. Biol. Chem. 265:6104-6111(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87162490; PubMed=3030813;			
RA	Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;			
RT	"Molecular cloning and characterization of a full-length cDNA clone for human plasminogen.";			
RL	FEBS Lett. 213:254-260(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS LYS-57; GLN-133; HIS-261; TRP-408;			
RP	ASN-472; VAL-494 AND TRP-523.			
RA	Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,			
RA	Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;			
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE OF 20-810, AND VARIANT ASN-472.			
RA	Sottrup-Jensen L., Petersen T.E., Magnusson S.;			
RL	Submitted (JUL-1977) to the PIR data bank.			
RN	[5]			
RP	SEQUENCE OF 292-810 FROM N.A.			
RX	MEDLINE=85023311; PubMed=6148961;			
RA	Malinowski D.P., Sadler J.E., Davie E.W.;			
RT	"Characterization of a complementary deoxyribonucleic acid coding for human and bovine plasminogen.";			
RL	Biochemistry 23:4243-4250(1984).			
RN	[6]			
RP	SEQUENCE OF 20-100.			
RX	MEDLINE=75093329; PubMed=122932;			
RA	Wiman B., Wallen P.;			
RT	"Structural relationship between 'glutamic acid' and 'lysine' forms of human plasminogen and their interaction with the NH2-terminal activation peptide as studied by affinity chromatography.";			
RL	Eur. J. Biochem. 50:489-494(1975).			
RN	[7]			
RP	SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810, AND VARIANT ASN-472.			
RA	Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;			
RL	(In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);			

34	405.5	36.4	247	1	TRY2_CANFA	P06872 canis famil
35	405.5	36.4	490	1	TMS2_MOUSE	Q9j1q8 mus musculu
36	404	36.3	638	1	KAL_MOUSE	P26262 mus musculu
37	402.5	36.2	243	1	TRY1_BOVIN	P00760 bos taurus
38	402.5	36.2	246	1	TRY1_RAT	P00762 rattus norv
39	401	36.0	314	1	TEST_HUMAN	Q9y6m0 homo sapien
40	401	36.0	603	1	FA12_CAVPO	Q04962 cavia porce
41	401	36.0	638	1	KAL_RAT	P14272 rattus norv
42	400	35.9	270	1	TRYT_MERUN	P50342 meriones un
43	400	35.9	855	1	ST14_HUMAN	Q9y5y6 homo sapien
44	398.5	35.8	455	1	TMS5_MOUSE	Q9er04 mus musculu
45	398	35.8	436	1	ACRO_MOUSE	P23578 mus musculu

ALIGNMENTS

- RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
PL Raven Press, New York (1978).
RN [8]
- RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
PL plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [9]
- RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [10]
- RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
PL of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [11]
- RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
PL sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [12]
- RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Valli Z., Pattrly L.;
RT "Structure of the omega-aminocarboxylic acid-binding sites of human
PL plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [13]
- RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vali Z., Pattrly L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
PL are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [14]
- RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
PL plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [15]
- RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RT Gerwig G.J., van Halbeek H., Vliegthart J.F.;
PL "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to sialylation and
PL fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [16]
- RP CARBOHYDRATE-LINKAGE SITE SER-268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RT Pizzo S.V.;
PL "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
RT human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [17]
- RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
PL Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
PL suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [18]
- RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RT Lapcevic R., Nacy C.A.;
PL "A recombinant human angiostatin protein inhibits experimental primary
RT and metastatic cancer.";
RL Cancer Res. 57:1329-1334(1997).
RN [19]
- RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
PL refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [20]
- RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
PL human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [21]
- RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RX Sec B., Yamano A., Whitlow M., Teeter M.M.;
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
PL A possible structural role of disordered residues.";
RL Acta Crystallogr. D 53:169-178(1997).
RN [22]
- RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
RX MEDLINE=96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
PL plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [23]
- RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RT Castellino F.J.;
PL "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [24]
- RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
PL kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [25]
- RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
PL human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [26]
- RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194156; PubMed=8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RT Rickli E.E.;
PL "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
PL of plasminogen kringle domains";
RL Biochemistry 35:2357-2364(1996).
RN [27]
- RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;


```

FT DISULFID 699 766 BY SIMILARITY.
FT DISULFID 729 745 BY SIMILARITY.
FT DISULFID 756 784 BY SIMILARITY.
FT CARBOHYD 365 365 O-LINKED (GALNAC... ) (BY SIMILARITY).
SQ SEQUENCE 810 AA; 90255 MW; A75E1C51A1A0F24A CRC64;

Query Match 94.1%; Score 1047; DB 1; Length 810;
Best Local Similarity 92.8%; Pred. No. 4.8e-93;
Matches 192; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 MHFCGGTLISPEWVLTAAHCLSKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60
DB 604 MHFCGGTLISPEWVLTAAHCLSKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 663
QY 61 KOTALKLSPPAVITDKVIPACLPSPNVVADRTCEFTGWTGQTFGAGLKEAQLPV 120
DB 664 ADIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTGWTGQTFGAGLKEAQLPV 723
QY 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
DB 724 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 783
QY 181 CARPKGCVYVRVSRFVTWIEGVNRNN 207
DB 784 CARPKGCVYVRVSRFVTWIEGVNRNN 810

RESULT 3
ID PLMN CANFA STANDARD; PRT; 333 AA.
AC P80009;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RC MEDLINE=90175323; PubMed=2626424;
RX Schaller J., Straub C., Kaempfer U., Rickli E.E.;
RT "Complete amino acid sequence of canine miniplasminogen.";
RL Protein Seq. Data Anal. 2:445-450(1989).
CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion, and
CC inflammation; in ovulation it weakens the walls of the Graafian
CC follicle. It activates the urokinase-type plasminogen activator,
CC collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
CC Willebrand factor.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Activated with urokinase and high concentrations of
CC streptokinase.
CC -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot.
CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -1- SIMILARITY: Contains at least 1 kringle domain.
DR HSPG; P00747; 5HPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR0059003; Cys_Ser_trypsin.
DR InterPro; IPR0000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.

```

```

DR PFam; PF00051; kringle; 1.
DR PFam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00118; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PRODOM; PD000395; kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT NON_TER 1
FT CHAIN <1 103 PLASMIN HEAVY CHAIN A.
FT CHAIN 104 333 PLASMIN LIGHT CHAIN B.
FT DOMAIN 4 83 KRINGLE 5.
FT DOMAIN 104 333 SERINE PROTEASE.
FT DISULFID 4 83 BY SIMILARITY.
FT DISULFID 25 66 BY SIMILARITY.
FT DISULFID 54 78 BY SIMILARITY.
FT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).
FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).
FT DISULFID 130 146 BY SIMILARITY.
FT DISULFID 222 289 BY SIMILARITY.
FT DISULFID 252 268 BY SIMILARITY.
FT DISULFID 279 307 BY SIMILARITY.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;

Query Match 91.5%; Score 1018; DB 1; Length 333;
Best Local Similarity 91.3%; Pred. No. 1e-90; Mismatches 8; Indels 0; Gaps 0;
Matches 189; Conservative 10;

QY 2 HFCGGTLISPEWVLTAAHCLSKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 61
DB 128 HFCGGTLISPEWVLTAAHCLSKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 187
QY 62 DIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTGWTGQTFGAGLKEAQLPV 121
DB 188 DIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTGWTGQTFGAGLKEAQLPV 247
QY 122 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 181
DB 248 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 307
QY 182 ARPKGCVYVRVSRFVTWIEGVNRNN 207
DB 308 ARPKGCVYVRVSRFVTWIEGVNRNN 333

RESULT 4
ID APOA MACMU STANDARD; PRT; 1420 AA.
AC P14417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.

```

OK NCBI_TaxID=9544;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89174660; PubMed=2925643;
 RA Tomlinson J.E., McLean J.W., Lawn R.M.;
 RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
 RT synthesis";
 RL J. Biol. Chem. 264:5957-5965(1989).
 CC -1- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
 CC (LP(a)). It has serine proteinase activity and is able of
 CC autoproteolysis. Inhibits tissue-type plasminogen activator 1.
 CC Lp(a) may be a ligand for megalin/Gp 330.
 CC -1- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
 CC decorin (By similarity).
 CC -1- PTM: N- and O-glycosylated (By similarity).
 CC -1- DISEASE: Elevated plasma concentrations of apo(a) and its
 CC naturally occurring proteolytic fragments are correlated with
 CC atherosclerosis. Homology with plasminogen kringle IV and V is
 CC thought to underlie the atherogenicity of the protein, because the
 CC fragments are competing with plasminogen for fibrinogen binding.
 CC -1- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
 CC leading to the formation of the so called mini-Lp(a). Apo(a)
 CC fragments accumulate in atherosclerotic lesions, where they may
 CC promote thrombogenesis. O-glycosylation may limit the extent of
 CC proteolytic fragmentation (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 CC -1- SIMILARITY: Contains at least 10 kringle domains.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J04635; AAA36833.1; -;
 DR PIR; A32869; A32869.
 DR HSP; P00747; 2PKA.
 DR MEROPS; S01.226; -;
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 10.
 DR SMART; SM00130; KR; 10.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 10.
 DR PROSITE; PS00070; KRINGLE_2; 10.
 DR PROSITE; PS00240; TRYPIN_DOM; 1.
 DR PROSITE; PS00134; TRYPIN_HIS; FALSE NEG.
 DR PROSITE; PS00135; TRYPIN_SER; FALSE NEG.
 DR Hydrolase; Serine protease; Lipid transport; plasma; Glycoprotein;
 KW Kringle; Repeat; Atherosclerosis.
 FT NON_TER 1
 FT DOMAIN 1 49 127 KRINGLE 1.
 FT DOMAIN 163 241 KRINGLE 2.
 FT DOMAIN 277 355 KRINGLE 3.
 FT DOMAIN 391 469 KRINGLE 4.
 FT DOMAIN 505 583 KRINGLE 5.
 FT DOMAIN 619 697 KRINGLE 6.
 FT DOMAIN 725 803 KRINGLE 7.
 FT DOMAIN 839 917 KRINGLE 8.
 FT DOMAIN 953 1031 KRINGLE 9.
 FT DOMAIN 1067 1145 KRINGLE 10.
 FT DOMAIN 1191 1420 SERINE PROTEASE.
 SQ SEQUENCE 1420 AA; 158367 MW; BE10294903CSOE CRC64;
 Query Match 85.4%; Score 950; DB 1; Length 1420;

Best Local Similarity 85.9%; Pred. No. 2.2e-83;
 Matches 177; Conservative 9; Mismatches 20; Indels 0; Gaps 0;
 QY 2 HFCGGTLLISPEWVLTAAHCLKSPSSYKVLGAHQEVNLEPHVQIEVSRLEPTRK 61
 DB 1215 HFCGGTLLISPEWVLTAAHCLKSPSSYKVLGAHQEVNLEPHVQIEVSRLEPTRK 1274
 QY 62 DIALLKLSPPAVITDKVIPACLPSPNTVADRTCEFTGTGAGLKEAQLPVI 121
 DB 1275 DIALLKLSPPAVITDKVIPACLPSPNTVADRTCEFTGTGAGLKEAQLPVI 1334
 QY 122 ENKVCNRYEFLNGRVOSTELCAGHAGTSCQSDGGPLVCPEKDKVILQGVTSWGLGC 181
 DB 1335 ENTVCNHYEFLNGRVOSTELCAGHAGTSCQSDGGPLVCPEKDKVILQGVTSWGLGC 1394
 QY 182 ARNPKPGVYVVRVSVFTWIEGVWRNN 207
 DB 1395 ACNPKPGVYVVRVSVFTWIEGVWRNN 1420
 RESULT 5
 PLMN_PIG STANDARD; PRT; 790 AA.
 ID PLMN_PIG STANDARD; PRT; 790 AA.
 AC P06867;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7).
 GN PLG.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE OF 1-560.
 RA Schaller J., Marti T., Roessle S.J., Kaempfer U., Rickli E.E.;
 RT "Amino acid sequence of the heavy chain of porcine plasmin. Comparison
 RT of the carbohydrate attachment sites with the human and bovine
 RT species";
 RL Fibrinolysis 1:91-102(1987).
 RN [2]
 RP SEQUENCE OF 450-790.
 RX MEDLINE=85203907; PubMed=3846533;
 RA Marti T., Schaller J., Rickli E.E.;
 RT "Determination of the complete amino-acid sequence of porcine
 RT miniplasminogen";
 RL Eur. J. Biochem. 149:279-285(1985).
 RN [3]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 RT plasminogen. Species specificity in relation to sialylation and
 RT fucosylation patterns";
 RL Eur. J. Biochem. 173:57-63(1988).
 CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
 CC a proteolytic factor in a variety of other processes including
 CC embryonic development, tissue remodeling, tumor invasion, and
 CC inflammation; in ovulation it weakens the walls of the Graafian
 CC follicle. It activates the urokinase-type plasminogen activator,
 CC collagenases and several complement zymogens, such as C1 and C5.
 CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
 CC Willebrand factor.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen
 CC activators, both plasminogen and its activator being bound to
 CC fibrin. Cannot be activated with streptokinase.
 CC -1- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND
 CC IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAc
 CC DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES

```

CC CC (MICROHETEROGENEITY).
CC CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC CC immediately after dissociation from the clot.
CC CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC CC -!- SIMILARITY: Contains 5 kringle domains.
CC CC PIR; S03733; PLFG.
CC CC HSP; P00747; SHPG.
CC CC MEROPS; S01.233; -.
CC CC GlycoSuiteDB; P06867; -.
CC CC InterPro; IPR009003; Cys_Ser_trypsin.
CC CC InterPro; IPR000001; Kringle.
CC CC InterPro; IPR003014; PAN.
CC CC InterPro; IPR003609; Pan_app.
CC CC InterPro; IPR001254; Peptidase_S1.
CC CC InterPro; IPR001314; Peptidase_S1A.
CC CC InterPro; IPR003966; Peptidase_S1A_pr.
CC CC Pfam; PF00051; kringle; 5.
CC CC Pfam; PF00024; PAN; 1.
CC CC Pfam; PF00089; trypsin; 1.
CC CC PRINTS; PR00722; CHYMOTRYPSIN.
CC CC PRINTS; PR00018; KRINGLE.
CC CC PRINTS; PR01505; PROTHROMBIN.
CC CC ProDom; PD000395; Kringle; 5.
CC CC SMART; SM00130; KR; 5.
CC CC SMART; SM00473; PAN AP; 1.
CC CC SMART; SM00020; TRYD SPc; 1.
CC CC PROSITE; PS00021; KRINGLE 1; 5.
CC CC PROSITE; PS50070; KRINGLE^2; 5.
CC CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC CC PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
CC CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC CC Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
CC CC Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
CC CC CHAIN 1 560 PLASMIN HEAVY CHAIN A.
CC CC CHAIN 561 790 PLASMIN LIGHT CHAIN B.
CC CC DOMAIN 561 790 SERINE PROTEASE.
CC CC DOMAIN 84 162 KRINGLE 1.
CC CC DOMAIN 166 243 KRINGLE 2.
CC CC DOMAIN 256 333 KRINGLE 3.
CC CC DOMAIN 358 435 KRINGLE 4.
CC CC DOMAIN 461 540 KRINGLE 5.
CC CC ACT SITE 602 602 CHARGE RELAY SYSTEM.
CC CC ACT SITE 645 645 CHARGE RELAY SYSTEM.
CC CC ACT SITE 740 740 CHARGE RELAY SYSTEM.
CC CC CARBOHYD 289 289 N-LINKED (GLCNAC. . .).
CC CC CARBOHYD 340 340 /FTID-CAR 000019.
CC CC CARBOHYD 340 340 O-LINKED (GALNAC. . .).
CC CC CARBOHYD 340 340 /FTID-CAR 000020.
CC CC SEQUENCE 790 AA; 88592 MW; F04EA06E74BCD58E CRC64;
CC CC Query Match 85.0%; Score 946; DB 1; Length 790;
CC CC Best Local Similarity 84.5%; Pred. No. 2.6e-83;
CC CC Matches 174; Conservative 13; Mismatches 19; Indels 0; Gaps 0;
QY 2 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTRK 61
DB 585 HFCGGTLLSPWVLTAKHCKLEKSSSSSYKVLGAHEHYHLGEGVQIDVSKLFKEPSEA 644
QY 62 DIALLKLSAPVITDKVIPACLPSPNVVADRTCEFTTGTGCTGTGAGLLKEAQLPVI 121
DB 645 DIALLKLSAPVITDKVIPACLPSPNVVADRTACVITGTGCTGTGAGLLKEAQLPVI 704
QY 122 ENKVCNRYEFLNGRVQSTELCAGHLAGTSCQDGGPLVCPKDKVILQGVTSWGLGC 181
DB 705 ENKVCNRYEFLNGRVQSTELCAGHLAGTSCQDGGPLVCPKDKVILQGVTSWGLGC 764
QY 182 ARPNKPGYTVRSRVFTWIEGVWENN 207
DB 765 ALPNKPGYTVRSRVFTWIEGVWENN 790

```

RESULT 6

APOA_HUMAN

```

ID APOA_HUMAN STANDARD; PRT; 4548 AA.
AC P08519;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
GN LPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88039109; PubMed=3670400;
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA Fless G.M., Scannu A.M., Lawn R.M.;
RT "cDNA sequence of human apolipoprotein(a) is homologous to
RT plasminogen.";
RL Nature 330:132-137 (1987).
RN [2]
RP SERINE PROTEASE ACTIVITY.
RX MEDLINE=90076123; PubMed=2531657;
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
RT activity capable of cleaving it.";
RL EMBO J. 8:4035-4040 (1989).
RN [3]
RP REVIEW.
RX MEDLINE=90049223; PubMed=2530631;
RA Utermann G.;
RT "The mysteries of lipoprotein(a).";
RL Science 246:904-910 (1989).
RN [4]
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX MEDLINE=21303595; PubMed=11294842;
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
RT "Structural elucidation of the N- and O-glycans of human
RT apolipoprotein(a): role of o-glycans in conferring protease
RT resistance.";
RL J. Biol. Chem. 276:22200-22208 (2001).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX MEDLINE=96217891; PubMed=8642595;
RA Mikol V., Lofrasso P.V., Boettcher B.R.;
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT complexed with 6-aminocaproic acid and with p-aminomethylbenzoic
RT acid: existence of novel and expected binding modes.";
RL J. Mol. Biol. 256:751-761 (1996).
RN [6]
RP VARIANT ARG-4193.
RX MEDLINE=95002201; PubMed=7918682;
RA Scannu A.M., Pfaffinger D., Lee J.C., Hinman J.;
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45 (1994).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autophoretolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Op 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuNAcalpha2-3Galbeta1-3GalNAc, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringles IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,

```

leading to the formation of the so called mini-Lp(a). Apo(a) fragments accumulate in atherosclerotic lesions, where they may promote thrombogenesis. O-glycosylation may limit the extent of proteolytic fragmentation.

-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.

-!- SIMILARITY: Contains 38 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X06290; CAA29618.1; .
 PIR; S00657; S00657.
 PDB; 1I71; 13-JUN-01.
 PDB; 1JFN; 28-JUN-02.
 PDB; 1KIV; 18-MAY-99.
 PDB; 3KIV; 18-MAY-99.
 PDB; 4KIV; 18-MAY-99.
 MEROPS; S01-226; .
 Genew; HGNC:6667; LPA.
 MIM; 152200; .
 GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
 GO; GO:0008015; P:circulation; TAS.
 GO; GO:0009405; P:pathogenesis; TAS.
 InterPro; IPR009003; Cys Ser trypsin.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Peptidase S1.
 InterPro; IPR001314; Peptidase_S1A.
 Pfam; PF00051; kringle; 38.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00018; KRINGLE.
 ProDom; PD000395; Kringle; 38.
 SMART; SM00130; KR; 38.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS00021; KRINGLE_1; 38.
 PROSITE; PS50070; KRINGLE_2; 38.
 PROSITE; PS50240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
 Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.

FT SIGNAL 1 19
 FT CHAIN 20 4548 APOLIPOPROTEIN(A).
 FT DOMAIN 20 130 KRINGLE TYPE IV, 1.
 FT DOMAIN 131 244 KRINGLE TYPE IV, 2.
 FT DOMAIN 245 358 KRINGLE TYPE IV, 3.
 FT DOMAIN 359 472 KRINGLE TYPE IV, 4.
 FT DOMAIN 473 586 KRINGLE TYPE IV, 5.
 FT DOMAIN 587 700 KRINGLE TYPE IV, 6.
 FT DOMAIN 701 814 KRINGLE TYPE IV, 7.
 FT DOMAIN 815 928 KRINGLE TYPE IV, 8.
 FT DOMAIN 929 1042 KRINGLE TYPE IV, 9.
 FT DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
 FT DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
 FT DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
 FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
 FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
 FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
 FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
 FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
 FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
 FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
 FT DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
 FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
 FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
 FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
 FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
 FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.

FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
 FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
 FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
 FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
 FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
 FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
 FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
 FT DOMAIN 3665 3770 KRINGLE TYPE IV, 33.
 FT DOMAIN 3771 3884 KRINGLE TYPE IV, 34.
 FT DOMAIN 3885 3998 KRINGLE TYPE IV, 35.
 FT DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
 FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
 FT DOMAIN 4227 4327 KRINGLE TYPE V.
 FT DOMAIN 4328 4548 SERINE PROTEASE.
 FT ACT_SITE 4329 4369 CHARGE RELAY SYSTEM.
 FT ACT_SITE 4412 4412 CHARGE RELAY SYSTEM.
 FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
 FT VARIANT 4193 4193 W -> R (LOSS OF LYSINE-SEPAPHAROSE BINDING).
 FT FT /FTID=VAR_006633.
 SQ SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;
 Query Match 83.9%; Score 933.5; DB 1; Length 4548;
 Best Local Similarity 84.5%; Pred. No. 3.6e-81;
 Matches 174; Conservative 11; Mismatches 12; Indels 9; Gaps 1;
 Qy 2 HFCGGLISPEWVLTAAHCLKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTRK 61
 Db 4352 HFCGGLISPEWVLTAAHCLKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTRK 4411
 Qy 62 DIALLLKSPAVITDKVIPACLPSPNYVADRTCEFTMGWGETQGTGAGLLKKAQLPVI 121
 Db 4412 DIALLLKSPAVITDKVPACLPSPDYMTARTCEYITMGWGETQGTGAGLLKKAQLPVI 4471
 Qy 122 ENKVCNRYEFLNGRVQSTELCAGHAGTSCQSDGGPLVCFPKDKYILQGVTSWGLGC 181
 Db 4472 ENEVCNHYK-----TCAEHLARTDSCQSDGGPLVCFPKDKYILQGVTSWGLGC 4522
 Qy 182 ARPKNPGVYVRVSFRFTWIEGVMNN 207
 Db 4523 ARPKNPGVYVRVSFRFTWIEGVMNN 4548
 RESULT 7
 ID PLMN MOUSE STANDARD; PRT; 812 AA.
 AC P20918; Q8CIS2; Q91WJ5;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
 GN PLG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91184812; PubMed=2081600;
 RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
 RT "Characterization of the cDNA coding for mouse plasminogen and
 RT localization of the gene to mouse chromosome 17.";
 RL Genomics 8:49-61(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
 RA Nagaraja R.;
 RT "Genomic sequence analysis in the mouse t-complex region.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;

RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,
RA Bobak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smallos D.E.,
RA Smerchek A., Schein J.B., Jones S.J.M., Matra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[4]
RP SEQUENCE OF 1-16 FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Liver;
RX PubMed=12149246;
RA Bannach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L.,
RA Parmer R.J., Miles L.A.,
RT "Localization of regulatory elements mediating constitutive and
RT cytokine-stimulated plasminogen gene expression.";
RL J. Biol. Chem. 277:38579-38588 (2002).
[5]
RN CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RP MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328 (1994).
CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion, and
CC inflammation; in ovulation it weakens the walls of the Graafian
CC follicle. It activates the urokinase-type plasminogen activator,
CC collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
CC Willebrand factor.
CC -1- FUNCTION: Angiostatin is an angiogenesis inhibitor that blocks
CC neovascularization and growth of experimental primary and
CC metastatic tumors in vivo.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Cannot be activated with streptokinase.
CC -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot.
CC -1- MISCELLANEOUS: In the presence of the inhibitor, the activation
CC involves only cleavage after Arg-581, resulting in 2 chains held
CC together by 2 disulfide bonds. Without the inhibitor, the
CC activation involves also removal of the activation peptide.
CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -1- SIMILARITY: Contains 5 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; J04766; AAA50168.1; --

DR EMBL; AF481053; AAM22156.1; --
DR EMBL; BC014773; AAH14773.1; --
DR EMBL; BC057186; AAH57186.1; --
DR EMBL; AY134430; AAN15805.1; --
DR PIR; A38514; PLMS.
DR HSP; P00747; IPMK.
DR MEROPS; S01.233; --
DR MGD; MGI:97620; Plg.
DR GO; GO:0016506; P:apoptosis activator activity; IDA.
DR GO; GO:0006915; P:apoptosis; IDA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 812 PLASMINOGEN.
FT CHAIN 20 581 PLASMIN HEAVY CHAIN A.
FT PEPTIDE 20 97 ACTIVATION PEPTIDE.
FT CHAIN 98 581 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 98 7436 ANGIOSTATIN.
FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT DOMAIN 582 812 SERINE PROTEASE.
FT ACT_SITE 624 624 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 568 687 INTERCHAIN (BY SIMILARITY).
FT DISULFID 578 586 INTERCHAIN (BY SIMILARITY).
FT DISULFID 609 625 BY SIMILARITY.
FT DISULFID 701 768 BY SIMILARITY.

FT DISULFID 731 747 BY SIMILARITY.
 FT DISULFID 758 786 BY SIMILARITY.
 FT CONFLICT 235 R -> H (IN REF. 1).
 FT CONFLICT 525 G -> D (IN REF. 1).
 FT CONFLICT 649 S -> L (IN REF. 1).
 SQ SEQUENCE 812 AA; 90781 MW; 241732606A2FFD2 CRC64;

Query Match 83.3%; Score 927; DB 1; Length 812;
 Best Local Similarity 84.0%; Pred. No. 1.8e-81;
 Matches 173; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

Qy 2 HFCGGTILSPFWLTAHCLKSPSSYKVLGAHQEVNLPVHVQIEVSRLEPTRK 61
 Db 607 HFCGGTILSPFWLTAHCLKSPSSYKVLGAHQEVNLPVHVQIEVSRLEPTRK 666

Qy 62 DIALLKLSPPAVITDKVIPACLPSPNPNVADRTCEFTGTCGAGLKEAQLPVI 121
 Db 667 DIALLKLSPPAVITDKVIPACLPSPNPNVADRTCEFTGTCGAGLKEAQLPVI 726

Qy 122 ENKVCNRYEFLNGRVQSTELCAGHAGTSCQDGGPLVCFPKDKYILQGVTSWGLGC 181
 Db 727 ENKVCNRYEFLNGRVQSTELCAGHAGTSCQDGGPLVCFPKDKYILQGVTSWGLGC 786

Qy 182 ARPNKPGVYVRSFVTVIEGWMM 207
 Db 787 ARPNKPGVYVRSFVTVIEGWMM 812

RESULT 8
 PLMN_SHEEP STANDARD; PRT; 343 AA.
 AC P81286;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragment).
 GN PLG.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93149995; PubMed=1492092;
 RA Schaller J., Straub C., Kampfer U., Rickli E.E.;
 RT "Complete amino acid sequence of ovine miniplasminogen.";
 RL Protein Seq. Data Anal. 5:21-25(1992).
 CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
 a proteolytic factor in a variety of other processes including
 embryonic development, tissue remodeling, tumor invasion, and
 inflammation; in ovulation it weakens the walls of the Graafian
 follicle. It activates the urokinase-type plasminogen activator,
 collagenases and several complement zymogens, such as C1 and C5.
 CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
 Willebrand factor.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 higher selectivity than trypsin. Converts fibrin into soluble
 products.
 CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen
 activators, both plasminogen and its activator being bound to
 fibrin. Cannot be activated with streptokinase.
 CC -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
 immediately after dissociation from the clot.
 CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 CC -1- SIMILARITY: Contains at least 2 kringle domains.
 DR PIR; B61545; B61545.
 DR HSSP; P00747; SHPG.
 DR MEROPS; S01.233; -.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR003966; Peptidase_S1A_pr.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; tryptsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00133; TRYPSIN_SER; 1.
 KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
 FT NON_TER 1 1
 FT DOMAIN <1 140 HEAVY CHAIN A.
 FT DOMAIN 141 >343 LIGHT CHAIN A.
 FT DOMAIN <1 17 KRINGLE 4.
 FT DOMAIN 41 120 KRINGLE 5.
 FT DOMAIN 114 341 SERINE PROTEASE.
 FT ACT_SITE 181 181 CHARGE RELAY SYSTEM.
 FT ACT_SITE 224 224 CHARGE RELAY SYSTEM.
 FT ACT_SITE 319 319 CHARGE RELAY SYSTEM.
 FT NON_TER 343 343
 SQ SEQUENCE 343 AA; 37662 MW; 8DF6EBA92D596EE0 CRC64;

Query Match 82.9%; Score 923; DB 1; Length 343;
 Best Local Similarity 84.8%; Pred. No. 1.5e-81;
 Matches 173; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

Qy 2 HFCGGTILSPFWLTAHCLKSPSSYKVLGAHQEVNLPVHVQIEVSRLEPTRK 61
 Db 138 HFCGGTILSPFWLTAHCLKSPSSYKVLGAHQEVNLPVHVQIEVSRLEPTRK 197

Qy 62 DIALLKLSPPAVITDKVIPACLPSPNPNVADRTCEFTGTCGAGLKEAQLPVI 121
 Db 198 DIALLKLSPPAVITDKVIPACLPSPNPNVADRTCEFTGTCGAGLKEAQLPVI 257

Qy 122 ENKVCNRYEFLNGRVQSTELCAGHAGTSCQDGGPLVCFPKDKYILQGVTSWGLGC 181
 Db 258 ENKVCNRYEFLNGRVQSTELCAGHAGTSCQDGGPLVCFPKDKYILQGVTSWGLGC 317

Qy 182 ARPNKPGVYVRSFVTVIEGWMM 205
 Db 318 ARPNKPGVYVRSFVTVIEGWMM 341

RESULT 9
 PLMN_HORSE STANDARD; PRT; 338 AA.
 AC P80010;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragment).
 GN PLG.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=92052077; PubMed=1946332;
 RA Schaller J., Straub C., Kampfer U., Rickli E.E.;
 RT "Complete amino acid sequence of equine miniplasminogen.";
 RL Protein Seq. Data Anal. 4:69-74(1991).
 CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
 a proteolytic factor in a variety of other processes including
 embryonic development, tissue remodeling, tumor invasion, and
 inflammation; in ovulation it weakens the walls of the Graafian

CC follicle. It activates the urokinase-type plasminogen activator,
 CC collagenases and several complement zymogens, such as C1 and C5.
 CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
 CC Willebrand factor.

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.

CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen
 CC activators, both plasminogen and its activator being bound to
 CC fibrin. Activated with catalytic amounts of streptokinase.
 CC -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
 CC immediately after dissociation from the clot.

CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
 CC -1- SIMILARITY: Contains at least 1 kringle domain.

DR PIR; A61545; A61545.

DR HSSP; P00747; SHPG.

DR MEROPS; S01.233; -.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000001; Kringle_1.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR003966; Peptidase_S1A_pr.

DR Pfam; PF00051; Kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR PRINTS; PR01505; PROTHROMBIN.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease; plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.

FT NON_TER 1 1

FT CHAIN <1 108 PLASMIN HEAVY CHAIN A.

FT CHAIN 109 338 PLASMIN LIGHT CHAIN B.

FT DOMAIN 9 88 KRINGLE 5.

FT DOMAIN 109 338 SERINE PROTEASE.

FT DISULFID 9 88 BY SIMILARITY.

FT DISULFID 30 71 BY SIMILARITY.

FT DISULFID 59 83 BY SIMILARITY.

FT DISULFID 95 213 INTERCHAIN (BY SIMILARITY).

FT DISULFID 105 113 INTERCHAIN (BY SIMILARITY).

FT DISULFID 135 151 BY SIMILARITY.

FT DISULFID 227 294 BY SIMILARITY.

FT DISULFID 257 273 BY SIMILARITY.

FT DISULFID 284 312 BY SIMILARITY.

FT ACT_SITE 150 150 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 193 193 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 288 288 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT SITE 157 157 STREPTOKINASE-BINDING SITE (PROBABLE).

FT SITE 191 191 STREPTOKINASE-BINDING SITE (PROBABLE).

FT SITE 269 269 STREPTOKINASE-BINDING SITE (PROBABLE).

FT SITE 282 282 SITE OF SUBSTRATE SPECIFICITY

FT SITE (BY SIMILARITY).

FT SEQUENCE 338 AA; 37132 MW; 8B9E5B5C5CDBE01 CRC64;

Query Match 81.3%; Score 905; DB 1; Length 338;
 Best Local Similarity 79.1%; Pred. No. 8.1e-80;
 Matches 163; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPWLVTAHCLKSPSSYKVIICAHQVNLPHVQIEVSRFLPEPTRK 61

Db 133 HFCGGTLLSPWLVTAHCLKSPSSYKVIICAHQVNLPHVQIEVSRFLPEPTRK 192

Qy 62 DIALLKLSAPVITDKVIPACLPSPNVVADRTECFITGMGTGTFGAGLLKEAQLPVI 121

Db 193 DIALLKLSAPVITDKVIPACLPSPNVVADRTECFITGMGTGTFGAGLLKEAQLPVI 252

Qy 122 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSQCGDGGPLVCFEKDKVILQGVTSWGLGC 181
 Db 253 ENKVCNRYEFLNGRVOSTELCAGHLVGGVDSQCGDGGPLVCFEKDKVILQGVTSWGLGC 312
 Qy 182 ARPNKPGVYVVRVSRFVTWIEGVNRNN 207
 Db 313 ARPNKPGVYVVRVSRFVTWIEGVNRNN 338

RESULT 10

PLMN ERIEU

ID PLMN ERIEU STANDARD; PRT; 810 AA.

AC Q29485;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Plasminogen precursor (EC 3.4.21.7).

GN PLG.

OS Erinaceus europaeus (Western European hedgehog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.

OX NCBI_TaxID=9365;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=96025778; PubMed=7592597;

RA Lawn R.M., Boommark N.W., Schwartz K., Lindahl G.E., Wade D.P.,

RA Byrne C.D., Fong K.J., Meer K., Pathy L.;

RT "The recurring evolution of lipoprotein(a). Insights from cloning of

RT hedgehog apolipoprotein(a).";

RL J. Biol. Chem. 270:24004-24009(1995).

RN [2]

RP REVISIONS.

RA Lawn R.M.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as

CC a proteolytic factor in a variety of other processes including

CC embryonic development, tissue remodeling, tumor invasion, and

CC inflammation; in ovulation it weakens the walls of the Graafian

CC follicle. It activates the urokinase-type plasminogen activator,

CC collagenases and several complement zymogens, such as C1 and C5.

CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von

CC Willebrand factor.

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;

CC higher selectivity than trypsin. Converts fibrin into soluble

CC products.

CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen

CC activators, both plasminogen and its activator being bound to

CC fibrin. Cannot be activated with streptokinase.

CC -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin

CC immediately after dissociation from the clot.

CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.

CC -1- SIMILARITY: Contains 5 kringle domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U33171; AAC48717.1; -.

DR PIR; I46260; I46260.

DR HSSP; P00747; 1PMK.

DR MEROPS; S01.233; -.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan_app.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR003966; Peptidase_S1A_pr.

```
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PRODOM; PD000395; kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00021; KRINGLE 1; 5.
DR PROSITE; PS00070; KRINGLE 2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 810 PLASMINOGEN.
FT CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).
FT CHAIN 20 582 PLASMIN LIGHT CHAIN B (BY SIMILARITY).
FT DOMAIN 583 810 SERINE PROTEASE.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 185 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 379 456 KRINGLE 4.
FT DOMAIN 482 561 KRINGLE 5.
FT ACT_SITE 622 622 CHARGE RELAY SYSTEM.
FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
FT CARBOHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 810 AA; 90902 MW; 8575780946017A16 CRC64;

Query Match 80.1%; Score 891; DB 1; Length 810;
Best Local Similarity 79.1%; Pred. No. 5.3e-78;
Matches 163; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

QY 2 HFCGGTILSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLPHPVQIEVSRLEPTRK 61
DB 605 HFCGGTILSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLPHPVQIEVSRLEPTRK 664
QY 62 DIALKLLSSPAVITDKVIPACLPSPNYVADRTSCFTTGWGETGTFGAGLLKEAQLPVI 121
DB 665 DIALKLLSSPAVITDKVIPACLPSPNYVADRTSCFTTGWGETGTFGAGLLKEAQLPVI 724
QY 122 ENKVCNRYEFLNGRVOSTELCAGHLAGTSCQDGGPLVCPFKDKVILQGVTSWGLGC 181
DB 725 ENKVCNRYEFLNGRVOSTELCAGHLAGTSCQDGGPLVCPFKDKVILQGVTSWGLGC 784
QY 182 ARPNKPGVYVRVSRFTWIEGVNRRN 207
DB 785 ARLTRPGVYVRVSRVSVLQDWMNN 810

RESULT 11
ID PLMN_BOVIN STANDARD; PRT; 812 AA.
AC P06868; Q28162;
DT 01-JAN-1998 (Rel. 06, Created)
DT 15-MAR-2004 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Berglund L., Andersen M.D., Petersen T.E.;
```

```
RT "Cloning and characterization of the bovine plasminogen cDNA.";
RN Int. Dairy J. 5:593-603(1995).
RL [2]
RX SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=85203906; PubMed=3846532;
RA Schaller J., Moser P.W., Danneberg-Muller G.A.K., Rosselet S.J.,
RA Kampfer U., Rickli E.E.;
RT "Complete amino acid sequence of bovine plasminogen. Comparison with
RT human plasminogen.";
RL Eur. J. Biochem. 149:267-278(1985).
RN [3]
RX SEQUENCE OF 706-812 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [4]
RX CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to sialylation and
RT fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
CC -1- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion,
CC and inflammation; in ovulation it weakens the walls of the
CC Graafian follicle. It activates the urokinase-type plasminogen
CC activator, collagenases and several complement zymogens, such
CC as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin,
CC laminin and von Willebrand factor.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Cannot be activated with streptokinase.
CC -1- PM: N-LINKED GLYCAN CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.
CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS
CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
CC -1- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot.
CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -1- SIMILARITY: Contains 5 kringle domains.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X79402; CAA55939.1; -.
DR EMBL; K02935; AAA30714.1; -.
DR PIR; S45046; PLBO.
DR HSP; P00747; 2PK4.
DR MEROPS; S01.233; -.
DR GlycoSuiteDB; P06868; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_SIA.
DR InterPro; IPR003966; Peptidase_SIA_pr.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
```

DR PRINTS; PRO0018; KRINGLE.
 DR PRINTS; PRO1505; BROTHROMBIN.
 DR PRODOM; PD000395; Kringle; 5.
 DR SMART; SM00130; KR; 5.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; TRYP-Spc; 1.
 DR PROSITE; PS00021; KRINGLE 1; 5.
 DR PROSITE; PS00070; KRINGLE 2; 5.
 DR PROSITE; PS00240; TRYPIN_DOM; 1.
 DR PROSITE; PS00134; TRYPIN_HIS; 1.
 DR PROSITE; PS00135; TRYPIN_SER; 1.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
 KW Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 812 PLASMINOGEN.
 FT CHAIN 27 812 PLASMIN HEAVY CHAIN A.
 FT CHAIN 27 812 PLASMIN LIGHT CHAIN B.
 FT CHAIN 584 812 KRINGLE 1.
 FT DOMAIN 110 188 KRINGLE 1.
 FT DOMAIN 192 269 KRINGLE 2.
 FT DOMAIN 282 359 KRINGLE 3.
 FT DOMAIN 384 461 KRINGLE 4.
 FT DOMAIN 485 564 KRINGLE 5.
 FT DOMAIN 584 812 SERINE PROTEASE.
 FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 315 315 /FTid=CAR_000014.
 FT CARBOHYD 365 365 O-LINKED (GALNAC. . .).
 FT CARBOHYD 365 365 /FTid=CAR_000015.
 FT ACT SITE 624 624 CHARGE RELAY SYSTEM.
 FT ACT SITE 667 667 CHARGE RELAY SYSTEM.
 FT ACT SITE 762 762 CHARGE RELAY SYSTEM.
 FT CONFLICT 335 335 N -> D (IN REF. 2).
 FT CONFLICT 516 516 Q -> H (IN REF. 2).
 FT CONFLICT 555 555 P -> L (IN REF. 2).
 FT CONFLICT 744 744 T -> R (IN REF. 3).
 SQ SEQUENCE 812 AA; 91216 MW; 38A6AA691E220946 CRC64;
 Query Match 78.9%; Score 878; DB 1; Length 812;
 Best Local Similarity 80.1%; Pred. No. 9.6e-77;
 Matches 165; Conservative 12; Mismatches 29; Indels 0; Gaps 0;
 QY 2 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHOEVNLEPHVQIEVSRFLFLEPTRK 61
 DB 607 HFCGGTLLSPKWLTAAHCLDNILALSFYKVLGAHNEKVQSQVQEIPLVSRFLFRESQA 666
 QY 62 DIALLLKSSPAVITDKVIPACLPSPNVVADRTSCFTTGHGETGTGAGLLKEAQLPVI 121
 DB 667 DIALLLKSRPAITKEVIPACLPSPNVVAAARTCYITGWGETGTGEGLLKEAHLFVI 726
 QY 122 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCQSGPLVCFKDKYILQGVTSWGLGC 181
 DB 727 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCQSGPLVCFKDKYILQGVTSWGLGC 786
 QY 182 ARPNKPGYVVRVSFVTWIEGWENN 207
 DB 787 ARPNKPGYVVRVSFVPMIETWERN 812
 RESULT 12
 TMS3 HUMAN
 ID TMS3 HUMAN STANDARD; PRT; 454 AA.
 AC P57727;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Transmembrane protease, serine 3 (EC 3.4.21.-) (Serine protease
 DE TDG-12) (Tumor associated differentially-expressed gene-12 protein).
 GN TMPSR3 OR TDG12 OR ECHOS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS A AND T).
 RC TISSUE-Ovarian carcinoma;
 RX MEDLINE=20521358; PubMed=11068177;
 RA Underwood L.J., Shigemasa K., Tanimoto H., Beard J.B., Schneider E.N.,
 RA Wang Y., Parmley T.H., O'Brien T.J.;
 RT "Ovarian tumor cells express a novel multi-domain cell surface serine
 RL protease.";
 RL Biochim. Biophys. Acta 1502:337-350(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS A; B AND D), AND VARIANT ILE-53.
 RX MEDLINE=20578749; PubMed=11137999;
 RA Scott H.S., Kudoh J., Wattenhofer M., Shibuya K., Berry A., Chraat R.,
 RA Guipponi M., Wang J., Kawasaki K., Asakawa S., Minoshima S.,
 RA Younus F., Mehdi S.Q., Radhakrishna U., Papasavvas M.P., Gehrig C.,
 RA Rossier C., Korostishevsky M., Gal A., Shimizu N., Bonne-Tamir B.,
 RA Antonarakis S.E.;
 RT "Insertion of beta-satellite repeats identifies a transmembrane
 RT protease causing both congenital and childhood onset autosomal
 RL recessive deafness.";
 RL Nat. Genet. 27:59-63(2001).
 RN [3]
 RP SUBCELLULAR LOCATION, AND FUNCTION IN ENAC CLEAVAGE.
 RX MEDLINE=22281255; PubMed=12393794;
 RA Guipponi M., Vuagniaux G., Wattenhofer M., Shibuya K., Vazquez M.,
 RA Dougherty L., Scamuffa N., Guida E., Okui M., Rossier C., Hancock M.,
 RA Buchet K., Raymond A., Hummler E., Marzella P.L., Kudoh J.,
 RA Shimizu N., Scott H.S., Antonarakis S.E., Rossier B.C.;
 RT "The transmembrane serine protease (TMPSR3) mutated in deafness
 RT DFNB8/10 activates the epithelial sodium channel (ENAC) in vitro.";
 RL Hum. Mol. Genet. 11:2829-2836(2002).
 RN [4]
 RP VARIANTS DFNB8/DFNB10 CYS-251 AND LEU-404.
 RX MEDLINE=21354482; PubMed=11462234;
 RA Masmoudi S., Antonarakis S.E., Schwede T., Ghorbel A.M., Gratri M.,
 RA Pappasavvas M.P., Drira M., Elgaied-Bouilla A., Wattenhofer M.,
 RA Rossier C., Scott H.S., Ayadi H., Guipponi M.;
 RT "Novel missense mutations of TMPSR3 in two consanguineous Tunisian
 RT families with non-syndromic autosomal recessive deafness.";
 RL Hum. Mutat. 18:101-108(2001).
 RN [5]
 RP VARIANTS DFNB8/DFNB10 TRP-109, PHE-194 AND ARG-407, AND VARIANTS
 RP ILE-53; SER-111 AND VAL-253.
 RX MEDLINE=21317610; PubMed=11424922;
 RA Ben-Yosef T., Wattenhofer M., Riazuddin S., Ahmed Z.M., Scott H.S.,
 RA Kudoh J., Shibuya K., Antonarakis S.E., Bonne-Tamir B.,
 RA Radhakrishna U., Naz S., Ahmed Z., Riazuddin S., Pandya A.,
 RA Nance W.E., Wilcox E.R., Friedman T.B., Morell R.J.;
 RT "Novel mutations of TMPSR3 in four DFNB8/B10 families segregating
 RT congenital autosomal recessive deafness.";
 RL J. Med. Genet. 38:396-400(2001).
 RN [6]
 RP VARIANT DFNB8/DFNB10 GLY-103, AND VARIANTS ASN-173 AND THR-426.
 RX MEDLINE=21904597; PubMed=11907649;
 RA Wattenhofer M., Di Iorio V., Rabionet R., Dougherty L., Papanos A.,
 RA Schwede T., Montserrat-Sentis B., Arbones L., Illades T.,
 RA Pasquidibiscaglia A., D'Amelio M., Alban S., Rossier C., Dahl H.-H.M.,
 RA Petersen M.B., Estivill X., Gasparini P., Scott H.S.,
 RA Antonarakis S.E.;
 RT "Mutations in the TMPSR3 gene are a rare cause of childhood
 RT nonsyndromic deafness in Caucasian patients.";
 RL J. Mol. Med. 80:124-131(2002).
 CC -1- FUNCTION: Probable protease. Seems to be capable of activating
 CC ENAC.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
 CC reticulum.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Evt=Alternative splicing; Named isoforms=4;
 CC Name=A;
 CC IsoId=P57727-1; Sequence=Displayed;
 CC Name=B; Synonyms=C;
 CC IsoId=P57727-2; Sequence=VSP_005391;
 CC Name=D;
 CC IsoId=P57727-3; Sequence=VSP_005392;
 CC

CC Name=T; Synonyms=Truncated, TADG-12V;
 CC IsoId=B57727-4; Sequence=VSP_005393, VSP_005394;
 CC -1- TISSUE SPECIFICITY: Expressed in many tissues including fetal
 CC cochlea. Isoform T is found at increased levels in some
 CC carcinomas.
 CC -1- PTM: Undergoes autophosphorylation.
 CC -1- DISEASE: Defects in TMPRSS3 are a cause of childhood-onset
 CC autosomal recessive neurosensory deafness 8 (DFNB8) [MIM:601072].
 CC -1- DISEASE: Defects in TMPRSS3 are a cause of congenital autosomal
 CC recessive neurosensory deafness 10 (DFNB10) [MIM:605316].
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -1- SIMILARITY: Contains 1 SRCR domain.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AF201380; AAC37012.1; -;
 CC EMBL: AB038157; BAB20077.1; -;
 CC EMBL: AB038158; BAB20078.1; -;
 CC EMBL: AB038159; BAB20079.1; -;
 CC EMBL: AB038160; BAB20080.1; -;
 CC HSSP: P00763; 1DPO.
 CC MEROPS: S01.079; -;
 CC Genew: HGNC:11877; TMPRSS3.
 CC MIM: 605511; -;
 CC MIM: 601072; -;
 CC MIM: 605316; -;
 CC GO: GO:0016021; C: integral to membrane; NAS.
 CC GO: GO:0004252; F: serine-type endopeptidase activity; NAS.
 CC GO: GO:0006508; P: proteolysis and peptidolysis; NAS.
 CC InterPro: IPR009003; Cys_Ser_trypsin.
 CC InterPro: IPR002172; LDL_receptor_A.
 CC InterPro: IPR001254; Peptidase_S1.
 CC InterPro: IPR001314; Peptidase_S1A.
 CC InterPro: IPR001190; Srcr_receptor.
 CC Pfam: PF00057; ldl_recept_a; 1.
 CC Pfam: PF00089; trypsin; 1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00192; LDLA; 1.
 CC SMART: SM00202; SR; 1.
 CC SMART: SM00020; Tryp_SPC; 1.
 CC PROSITE: PS01209; LDLRA_1; 1.
 CC PROSITE: PS00668; LDLRA_2; 1.
 CC PROSITE: PS00420; SRCR_1; FALSE_NEG.
 CC PROSITE: PS00287; SRCR_2; 1.
 CC PROSITE: PS02040; TRYPSIN_DOM; 1.
 CC PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC PROSITE: PS00133; TRYPSIN_SER; 1.
 CC K0 Hydroxylase; Serine protease; Transmembrane; Signal-anchor; Zymogen;
 CC K0 Endoplasmic reticulum; Deafness; Alternative splicing;
 CC K0 Disease mutation; Polymorphism.
 CC DOMAIN 1 48
 CC TRANSNEM 49 69
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC LDL-RECEPTOR CLASS A.
 CC SRCR.
 CC SERINE PROTEASE.
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC CLEAVAGE (POTENTIAL).
 CC SITE 216 217
 CC BY SIMILARITY.
 CC DISULFID 73 85
 CC BY SIMILARITY.
 CC DISULFID 79 98
 CC BY SIMILARITY.
 CC DISULFID 92 107
 CC BY SIMILARITY.
 CC DISULFID 129 194
 CC BY SIMILARITY.
 CC DISULFID 142 204
 CC BY SIMILARITY.

FT DISULFID 207 324 BY SIMILARITY.
 FT DISULFID 258 BY SIMILARITY.
 FT DISULFID 338 BY SIMILARITY.
 FT DISULFID 370 BY SIMILARITY.
 FT DISULFID 397 BY SIMILARITY.
 FT CARBOHYD 221 N-LINKED (GLCNAC... (POTENTIAL).
 FT VARSPLIC 1 Missing (in isoform B).
 FT VARSPLIC 318 /FTid=VSP_005393.
 FT VARSPLIC 454 EMTQPVCLPNSSENFPPDKVCTWGTSGWATDGDAGDASPVLN
 FT HAAPVCLPNSSENFPPDKVCTWGTSGWATDGDAGDASPVLN
 FT DSGGGLVQBRRLKLVGATSGFICAEVKNKPGVYTVTSF
 FT LDMHEQMRDLKT -> GTSGSLCGSAALPLFQEDLQLLI
 FT EAPL (in isoform D).
 FT /FTid=VSP_005392.
 FT DLYLPKSWTIQVGLVSLDNPAPSHLVEKVIYH -> EIVA
 FT PRERADRGRKLLKWRKPTMKGRPPSHS (in isoform
 FT T).
 FT /FTid=VSP_005393.
 FT Missing (in isoform T).
 FT /FTid=VSP_005394.
 FT V -> I.
 FT /FTid=VAR_010781.
 FT D -> G (in DFNB8/DFNB10).
 FT /FTid=VAR_013490.
 FT R -> W (in DFNB8/DFNB10).
 FT /FTid=VAR_013491.
 FT G -> S.
 FT /FTid=VAR_013492.
 FT D -> N.
 FT /FTid=VAR_013493.
 FT C -> F (in DFNB8/DFNB10).
 FT /FTid=VAR_013494.
 FT W -> C (in DFNB8/DFNB10).
 FT /FTid=VAR_011678.
 FT I -> V (in dbSNP:2839500).
 FT /FTid=VAR_013101.
 FT P -> L (in DFNB8/DFNB10).
 FT /FTid=VAR_011679.
 FT C -> R (in DFNB8/DFNB10).
 FT /FTid=VAR_013495.
 FT A -> T.
 Query Match 40.3%; Score 448; DB 1; Length 454;
 Best Local Similarity 42.4%; Pred. No. 1.4e-35;
 Matches 92; Conservative 35; Mismatches 74; Indels 16; Gaps 5;
 Qy 2 HFCGGLISPEWVLTAAHCLKSPRSSYKVLGAHQEVNL----BPHVQIEVSRLLFL 56
 Db 240 HLCGGSVITPLMIITAAHCVYDLYLPKSWTIQVGL---VSLDNPAPSHLVEKVIYHSKY 296
 Qy 57 EPTT--KDIALKLSSPAVITDKVIPACLPSPNYVVAADRTECFITGWTGTFGAG--- 111
 Db 297 KPRKLGNDIALMKLAGPLTFENMIQPVCLPNSSENFPPDKVCTWGTSGWATD--GAGDAS 354
 Qy 112 -LKBAQLPVIENKVCNRYEFLNGRVQSTELCAGHAGGTDSCGDSGGPLVCEKDKYI 170
 Db 355 PVLNHAAPVLIISNKNHRDVGIGIISPLKAGYLTTGGVDSCGDSGGPLVCEKDKYI 414
 Qy 171 LQGVTSGLGCAEPKPGVYVVRVSRFVTWIEGVNRN 207
 Db 415 LVGATSGFICAEVKNKPGVYVVRVSRFVTWIEGVNRN 451
 RESULT 13
 TMS3 MOUSE
 ID TMS3 MOUSE STANDARD; PRT; 453 AA.
 AC QSKIT0; Q8VDE0;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Transmembrane protease, serine 3 (EC 3.4.21.-).
 GN TMPRSS3.
 OS Mus musculus (Mouse).

RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences."
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[3]

RN CHARACTERIZATION.
 RX MEDLINE=91358502; PubMed=1885621;
 RA Teuji A., Torres-Rosado A., Arai T., le Beau M.M., Lemons R.S.,
 RA Chou S.H., Kurachi K.;
 RT "Hepsin, a cell membrane-associated protease. Characterization,
 RT tissue distribution, and gene localization."
 RL J. Biol. Chem. 266:16948-16953(1991).

[4]

RN CHARACTERIZATION.
 RX MEDLINE=8346233; PubMed=8346233;
 RA Torres-Rosado A., O'Shea K.S., Teuji A., Chou S.H., Kurachi K.;
 RT "Hepsin, a putative cell-surface serine protease, is required for
 RT mammalian cell growth."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187(1993).

CC -!- FUNCTION: Plays an essential role in cell growth and maintenance
 CC of cell morphology.

CC -!- SUBCELLULAR LOCATION: Type II membrane protein.

CC -!- TISSUE SPECIFICITY: Present in most tissues, with the highest
 CC level in liver.

CC -!- SIMILARITY: Belongs to peptidase family S1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC

DR EMBL; M18930; AAA36013.1; -

DR EMBL; X07732; CAA30558.1; -

DR EMBL; X07002; CAA30058.1; -

DR EMBL; BC025716; AH25716.1; -

DR PIR; S00845; S00845.

DR HSP; P00763; 1DPO.

DR MEROPS; S01.224; -

DR Genew; HGNC:5155; HPN.

DR MIM; 142440; -

DR GO; GO:0005887; C:integral to plasma membrane; TAS.

DR GO; GO:0008236; F:serine-type peptidase activity; TAS.

DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.

DR InterPro; IPR009003; Cys Ser. trypsin.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease; Transmembrane; Signal-anchor.

FT SERINE PROTEASE HEPSIN, NON-CATALYTIC

FT CHAIN (POTENTIAL).

FT CHAIN 1 162

FT CHAIN 163 417

FT CYTOPLASMIC (POTENTIAL).

FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)

FT TRANSMEM 18 44

FT DOMAIN 1 17

FT EXTRACELLULAR (POTENTIAL).

FT DOMAIN 45 417

FT SERINE PROTEASE.

FT ACT_SITE 163 417

FT CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 203 203

FT CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 257 257

FT CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 353 353

FT INTERCHAIN (BY SIMILARITY).

FT DISULFID 153 277

FT BY SIMILARITY.

FT DISULFID 188 204

FT BY SIMILARITY.

FT DISULFID 322 338

FT BY SIMILARITY.

FT DISULFID 349 381

FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 112 112

SQ SEQUENCE 417 AA; 45011 MW; B2086FF661E551D7 CRC64;

Query Match 39.9%; Score 444; DB 1; Length 417;
 Best Local Similarity 42.2%; Pred. No. 3.1e-35;
 Matches 92; Conservative 38; Mismatches 66; Indels 22; Gaps 7;

QY 2 HPCGGLISPEWLTAAHCL-EKSPRSSYKVLGAHQEVNLEPHVQIEVSRL----- 54
 DB 186 HLCGSLSGDWLTAAHCFPRNRVLSRVRFAVAQAS--PHGLQLGVQAVVYHG 243
 QY 55 --PLEPRK---DIALKLSSPAVITDKVIPACLPSPNYPVADRTCTFGTGTGTF 108
 DB 244 LPRFDPNSENSNDIALVHLSPLTEYIQVCLPAAGQALVDGKICTVTGNGNTQ-YY 302
 QY 109 G--AGLLKEAQLPVIEKNRYEFLNGRVQSTELCAGHLAGTSDSCGSGGSLVCFEK 166
 DB 303 GQAGVLQEARVPIISNDVCGADFYGNQIKPKMFCAGYPEGGIDACQDGGGPFVCEDS 362
 QY 167 ----DKVILQGVTSWGLGCRPNKPGVYVVRVSRVTWI 200
 DB 363 ISRTPRRLCGIVSWGTCALAQKPGVYTKVDFREWI 400

RESULT 15

PR27_HUMAN STANDARD; PRT; 290 AA.
 AC Q9BQR3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pancreasin precursor (EC 3.4.21.-) (Marapsin) (Channel-activating
 DE protease 2) (CAPH2).
 GN PRSS27 OR MPN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fortunato M., Dando P.M., Rawlings N.D., Barrett A.J.;
 RT "Cloning, sequencing and expression of marapsin, a human serine
 RT proteinase";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Okaze H., Hayashi A., Kozuma S., Saito T.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22439795; PubMed=12441343;
 RA Bhagwandin V.J., Hau L.W.-T., Mallen-St Clair J., Wolters P.J.,
 RA Caughey G.H.;
 RT "Structure and activity of human pancreasin, a novel tryptic serine
 RT peptidase expressed primarily by the pancreas";
 RL J. Biol. Chem. 278:3363-3371(2003).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in the pancreas.
 CC -!- PTM: N-glycosylated.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AJ306593; CAC35467.1; -

DR EMBL; AB056161; BAB85497.1; -

DR EMBL; AY030095; AAK38168.1; -

DR HSP; P00734; IUVS.

DR MEROPS; S01.074; -

DR Genew; HGNC:15475; PRSS27.


```

DR MIN; 608018; --
DR InterPro; IPR009003; Cys Ser tryptsin.
DR InterPro; IPR01254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_SIA.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.
FT SIGNAL 1 22
FT PROPEP 23 34
FT CHAIN 35 290
FT DOMAIN 35 277
FT ACT_SITE 75 75
FT ACT_SITE 124 124
FT ACT_SITE 229 229
FT DISULFID 60 76
FT DISULFID 158 235
FT DISULFID 191 214
FT DISULFID 225 253
FT CARBOHYD 55 55
FT CARBOHYD 79 79
SQ SEQUENCE 290 AA; 31940 MW; 67BDC93EC70BFF7B CRC64;

Query Match      39.0%; Score 434; DB 1; Length 290;
Best Local Similarity 38.4%; Pred. No. 1.8e-34;
Matches 84; Conservative 43; Mismatches 76; Indels 16; Gaps 4;

Qy 2 HFCGGTILSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPH-----VQEIIEVSRFL 56
Db 58 HFCGSLIAECQWLTAAHCFRNTSETSLYQLLGARQLVQPGPHAMVAVRQVESNELYQ 117

Qy 57 -EPTKDIALLKSSPAVITDKVIPACLPSPNYVADRTECFITGK---ETQGTFGAGL 112
Db 118 GTASSADVALVEAPVPFTNYILPCLPDPSPVIFETGMNCWVTGWSPEEDLLPEPRI 177

Qy 113 LKEAQLPVIEKVCN-----RYEFLNGRVQSTELCAGHLAGGTDSCQDGGGGLVCFPE 165
Db 178 LQKLAVPIIDTPKCNLLYSKDTETGYQPKTKNDMLCAGFEKGKDKAGDGGGGLVCLV 237

Qy 166 KDKYILQGVTSWGLGCAKPNKPGVYVVRVSRFTVWIEGVM 204
Db 238 GQSWLQAGVISWGGGCAKQNRPGVYIRVTAHNNWIRII 276

```

Search completed: September 10, 2004, 14:52:51
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 14:47:48 ; Search time 40 Seconds
(without alignments)
497.791 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113

Sequence: 1 MHFCGGTLISPEWVLTAAHC.....GVYVRSRFTVWIEGVNRNN 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: Pirl1.*

2: Pirl2.*

3: Pirl3.*

4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1113	100.0	810	1 PLHU	plasmin (EC 3.4.21
2	1047	94.1	810	2 B30848	plasmin (EC 3.4.21
3	950	85.4	1420	2 A32869	apolipoprotein(a)
4	946	85.0	790	1 PUPG	plasmin (EC 3.4.21
5	933.5	83.9	4548	1 S00657	apoptein(a) (EC
6	925	83.1	812	1 PLMS	plasmin (EC 3.4.21
7	923	82.9	460	2 B61545	plasmin (EC 3.4.21
8	905	81.3	455	2 A61545	plasmin (EC 3.4.21
9	891	80.1	810	2 I46260	plasmin (EC 3.4.21
10	878	78.9	812	1 PUBO	plasmin (EC 3.4.21
11	444	39.9	417	1 S00845	hepsin (EC 3.4.21.
12	423	38.0	416	1 S33777	hepsin (EC 3.4.21.
13	423	38.0	1035	1 A43090	enteropeptidase (E
14	419	37.6	273	2 A47246	trypsin (EC 3.4.21
15	418	37.6	1019	1 A56318	enteropeptidase (E
16	416	37.4	1034	1 A53663	enteropeptidase (E
17	414	37.2	437	2 S18407	acrosin (EC 3.4.21
18	410.5	36.9	1524	2 T30337	polyprotein - Afri
19	409.5	36.8	231	1 TRPGTR	trypsin (EC 3.4.21
20	406	36.5	415	1 A34170	acrosin (EC 3.4.21
21	405.5	36.4	247	1 TRDG	trypsin (EC 3.4.21
22	404	36.3	638	1 KQMSPL	plasma kallikrein
23	402.5	36.2	229	1 TRBOT1	trypsin (EC 3.4.21
24	402.5	36.2	246	1 TRRT1	trypsin (EC 3.4.21
25	401	36.0	603	2 S28941	coagulation factor
26	401	36.0	638	1 KQRTPL	plasma kallikrein
27	400	35.9	270	2 S56160	mast cell trypsinase
28	398.5	35.8	418	2 A37344	acrosin (EC 3.4.21
29	398	35.8	436	2 JX0172	acrosin (EC 3.4.21

RESULT 1

PLHU

Plasmin (EC 3.4.21.7) precursor [validated] - human

N;Alternate names: plasminogen precursor [misnomer]

N;Contains: angiotatin; microplasmin; plasminogen

C;Species: Homo sapiens (man)

C;Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000

C;Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625; A04

R;Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.

J. Biol. Chem. 265, 6104-6111, 1990

A;Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr

A;Reference number: A35229; MUID:90202879; PMID:2318848

A;Accession: A35229

A;Molecule type: DNA

A;Residues: 1-810 <PRT>

A;Cross-references: GB:J05286; GB:M34276; NID:G190064; PIDN:AAA60113.1; PID:G387026

A;Experimental source: leukocyte; lung fibroblast

R;Malgarretti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; Ta

Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990

A;Title: Definition of the transcription initiation site of human plasminogen gene in 11

A;Reference number: I52242; MUID:91097523; PMID:2268308

A;Accession: I52242

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-16 <MAL1>

A;Cross-references: GB:M62890; NID:G190092; PIDN:AAA36454.1; PID:G553613

R;Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.

FEBS Lett. 213, 254-260, 1987

A;Title: Molecular cloning and characterization of a full-length cDNA clone for human pl

A;Reference number: A26646; MUID:87162490; PMID:3030813

A;Accession: A26646

A;Molecule type: mRNA

A;Residues: 1-471, 'D', 473-810 <FOR>

A;Cross-references: GB:X05199; NID:G35530; PIDN:CAA28831.1; PID:G35531

A;Experimental source: liver

R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.

Biochemistry 23, 4243-4250, 1984

A;Title: Characterization of a complementary deoxyribonucleic acid coding for human and

A;Reference number: I45961; MUID:85023311; PMID:6148961

A;Accession: I62738

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 292-471, 'D', 473-810 <MAL2>

A;Cross-references: GB:K02922; NID:G190112; PIDN:AAA60124.1; PID:G387031

A;Accession: I84609

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 367-419 <MAL3>

A;Cross-references: GB:K02921; NID:G190110; PIDN:AAA60123.1; PID:G190111

R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;

Eur. J. Biochem. 114, 465-470, 1981

A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,

A;Reference number: S03735; MUID:81212097; PMID:7238497
A;Accession: S03735
A;Molecule type: protein
A;Residues: 20-71,'E',73-76 <BRU>
R;Sottrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A;Reference number: A00929
A;Accession: A00929
A;Molecule type: protein
A;Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>
R;Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A;Title: Primary structure of the B-chain of human plasmin.
A;Reference number: A04627; MUID:7725245; PMID:142009
A;Accession: A04627
A;Molecule type: protein
A;Residues: 581-810 <WII>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A;Title: Structural relationship between "glutamic acid" and "lysine" forms of human plasminogen
A;Reference number: A04625; MUID:75093329; PMID:122932
A;Accession: A04625
A;Molecule type: protein
A;Residues: 20-50,'Q',51-71,'E',73-85,87-100 <W12>
R;Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A;Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen
A;Reference number: A04626; MUID:76043692; PMID:126863
A;Accession: A04626
A;Molecule type: protein
A;Residues: 483-507,'E',509-604 <W13>
R;Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A;Title: The primary structure of human plasminogen. II. The histidine loop of human plasminogen
A;Reference number: A92125; MUID:73149248; PMID:4694729
A;Contents: annotation; active site
R;Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A;Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A;Reference number: A92048; MUID:69234739; PMID:4240117
A;Contents: annotation; active site
R;Trexler, M.; Vali, Z.; Pathy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A;Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A;Reference number: A92382; MUID:82213905; PMID:6919539
A;Contents: annotation; omega-aminocarboxylic acid binding sites
R;Vali, Z.; Pathy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A;Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A;Reference number: A92458; MUID:85054794; PMID:6094526
A;Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R;Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A;Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative
A;Reference number: A58811; MUID:97067211; PMID:8910613
A;Contents: annotation
R;Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A;Title: Generation of an angiotensin-like fragment from plasminogen by streptolysin-1 (M
A;Reference number: A58812; MUID:9548733; PMID:9548733
A;Contents: annotation
R;Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51341; PDB:1PK4
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R;Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51488; PDB:2PK4
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R;Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A;Reference number: A51911; PDB:1PKR

A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R;Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A;Reference number: A52408; PDB:1PMK
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65244; PDB:1CEA
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65245; PDB:1CEB
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
A;Reference number: A58819; MUID:92031502; PMID:1657148
A;Contents: annotation
R;Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plasminogen
A;Reference number: A58818; MUID:92031503; PMID:1657149
A;Contents: annotation
R;de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.I.
Biochemistry 31, 270-279, 1992
A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.4 Å
A;Reference number: A39483; MUID:92118803; PMID:1310033
A;Contents: annotation; X-ray crystallography, 2.4 angstroms
R;Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A;Reference number: A65980; PDB:1KRN
A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R;Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65803; PDB:1HPJ
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65804; PDB:1HPK
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejzante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A;Reference number: S43645; MUID:94237157; PMID:8181475
A;Contents: annotation; conformation by (1)H-NMR, residues 96-184
R;Rejzante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
A;Reference number: A58817; MUID:94237158; PMID:8181476
A;Contents: annotation; conformation by (1)H-NMR
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other
C;Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU ar
C;Comment: Plasminogen is inactivated by alpha-2-antiplasmin (see PIR:THUA2) immediately after
C;Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial condit
C;Comment: Streptolysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiotensin. To
C;Genetics:
A;Gene: GDB:PIG
A;Cross-references: GDB:119498; OMIM:173350
A;Map position: 6q26-6q27
A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 529/
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a vari
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen acti
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology,
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydr
F;1-19/Domain: plasminogen-related protein precursor homology <PIPH>
F;1-96/Domain: signal sequence #status predicted <SIG>
F;20-810/Product: plasminogen #status experimental <PRO>
F;20-96/Domain: activation peptide #status experimental <APT>

F:79-466/Product: angiotensin #status experimental <AST>
F:97-580,581-810/Product: plasmin #status experimental <MAT>
F:97-580,Domain: plasmin chain A #status experimental <CHA>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 100.0%; Score 1113; DB 1; Length 810;
Best Local Similarity 100.0%; Pred. No. 1.5e-95; Indels 0; Gaps 0;
Matches 207; Conservative 0; Mismatches 0;

Qy 1 MHFCGGTLLSPWVLTAAHCKLEKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 604 MHFCGGTLLSPWVLTAAHCKLEKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 663
Qy 61 KDIALLLSSPAVITDKVIPACLPSPNVVADRTCEFTTGWGETQGTGAGLLKEAQLPV 120
Db 664 KDIALLLSSPAVITDKVIPACLPSPNVVADRTCEFTTGWGETQGTGAGLLKEAQLPV 723
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFPEKDKYILQGVTSWGLG 180
Db 724 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFPEKDKYILQGVTSWGLG 783

Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 784 CARPNKPGVYVRVSRFVTWIEGVNRN 810

RESULT 2

B30848
plasmin (EC 3.4.21.7) precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
C:Accession: B32869; B30848
R:Tominson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660; PMID:2925643
A:Accession: B32869
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-810 <TM>
A:Cross-references: GB:J04697; NID:G342272; PID:AAA36901.1; PID:G342273
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 324-352, 665, 760/Active site: His, Asp, Ser #status predicted

Query Match 94.1%; Score 1047; DB 2; Length 810;
Best Local Similarity 92.8%; Pred. No. 2.1e-89;
Matches 192; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPWVLTAAHCKLEKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 604 MHFCGGTLLSPWVLTAAHCKLEKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 663
Qy 61 KDIALLLSSPAVITDKVIPACLPSPNVVADRTCEFTTGWGETQGTGAGLLKEAQLPV 120
Db 664 ADIALLLSPAITDKVIPACLPSPNVVADRTCEFTTGWGETQGTGAGLLKEAQLPV 723
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFPEKDKYILQGVTSWGLG 180

Db 724 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFPEKDKYILQGVTSWGLG 783
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 784 CARPNKPGVYVRVSRFVTWIEGVNRN 810

RESULT 3

A32869
apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
C:Accession: A32869; A30848
R:Tominson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660; PMID:2925643
A:Accession: A32869
A:Molecule type: mRNA
A:Residues: 1-1420 <TM>
A:Cross-references: GB:J04635; NID:G342072; PID:AAA36833.1; PID:G342073
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:50-127/Domain: kringle homology <KR1>
F:164-241/Domain: kringle homology <KR2>
F:278-355/Domain: kringle homology <KR3>
F:392-469/Domain: kringle homology <KR4>
F:506-583/Domain: kringle homology <KR5>
F:620-697/Domain: kringle homology <KR6>
F:726-803/Domain: kringle homology <KR7>
F:840-917/Domain: kringle homology <KR8>
F:954-1031/Domain: kringle homology <KR9>
F:1068-1145/Domain: kringle homology <KR10>
F:1191-1413/Domain: trypsin homology <TRY>

Query Match 85.4%; Score 950; DB 2; Length 1420;
Best Local Similarity 85.9%; Pred. No. 4.5e-80;
Matches 177; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 61
Db 1215 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 1274
Qy 62 DIALLKLLSPAVITDKVIPACLPSPNVVADRTCEFTTGWGETQGTGAGLLKEAQLPV 121
Db 1275 DIALLKLLSPAVITDKVIPACLPSPNVVADRTCEFTTGWGETQGTGAGLLKEAQLPV 1334
Qy 122 ENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFPEKDKYILQGVTSWGLG 181
Db 1335 ENTVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFPEKDKYILQGVTSWGLG 1394
Qy 182 ARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 1395 ACNPKPGVYVRVSRFVTWIEGVNRN 1420

RESULT 4

PLPG
plasmin (EC 3.4.21.7) precursor - pig (fragment)

N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Sep-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C:Accession: S03733; S03737; A25834
R:Schaller, J.; Marti, T.; Roesele, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the ca
A:Reference number: S03733
A:Accession: S03733
A:Molecule type: protein
A:Residues: 1-560 <SCH>
R:Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;

Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human, A;Reference number: S03735; MUID:81212097; PMID:7238497
A;Accession: S03737
A;Molecule type: protein
A;Residues: 1-57 <BRU>
R;Martí, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A;Title: Determination of the complete amino-acid sequence of porcine miniplasminogen. A;Reference number: A25834; MUID:85203907; PMID:3846533
A;Accession: A25834
A;Molecule type: protein
A;Residues: 450-790 <MAR>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of the graafian follicle; also activates the urokinase-type plasminogen activator
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine protease
F;1-790/Product: plasminogen #status predicted <PRO>
F;1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>
F;1-77/Domain: activation peptide #status predicted <ACT>
F;78-560/Product: plasmin chain A #status predicted <AP>
F;84-162/Domain: kringle homology <KR1>
F;166-243/Domain: kringle homology <KR2>
F;256-333/Domain: kringle homology <KR3>
F;358-435/Domain: kringle homology <KR4>
F;450-790/Product: miniplasminogen #status experimental <MIN>
F;461-540/Domain: kringle homology <KR5>
F;561-790/Product: plasmin chain B #status experimental <BCH>
F;561-783/Domain: trypsin homology <TRY>
F;30-54, 34-42, 84-162, 105-145, 133-157, 166-243, 169-297, 187-226, 215-238, 256-333, 277-316, 305 bonds: #status predicted
F;602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 85.0%; Score 946; DB 1; Length 790;
Best Local Similarity 84.5%; Pred. No. 5.2e-80;
Matches 174; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

Qy 2 HFCGGTILSPDWLTAHCHLEKSPSSSYKVIILGAHQEVNLEPHVQIEVSRFLFETPRK 61
Db 585 HFCGGTILSPDWLTAHCHLEKSSPSYKVIILGAHQEVNLEPHVQIEVSRFLFETPRK 644

Qy 62 DIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTTGCTGCTGAGLLKEAQLPVI 121
Db 645 DIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTTGCTGCTGAGLLKEAQLPVI 704

Qy 122 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCQSGGSLVCFEKKYILQGVTSWGLGC 181
Db 705 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCQSGGSLVCFEKKYILQGVTSWGLGC 764

Qy 182 ARPNKPGVYVRSFVTWIEGVRNN 207
Db 765 ALPNKPGVYVRSFVTWIEGVRNN 790

RESULT 5
S00657
N;Alternate names: apolipoprotein(a) - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 08-Dec-2000
C;Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286
R;McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Scarlata, N.A.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.
Nature 330, 132-137, 1987
A;Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. A;Reference number: S00657; MUID:88039109; PMID:3670400
A;Accession: S00657
A;Molecule type: mRNA
A;Residues: 1-4548 <MCL>
A;Cross-references: GB:X06290; EMBL:X06696; NID:g28619; PIDN:CAN29618.1; PID:g28620
R;Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987

A;Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to A;Reference number: A28017; MUID:87204109; PMID:3472206
A;Accession: A28017
A;Molecule type: protein
A;Residues: 20-21, 'P', 23-34, 177-179, 'N', 181-186, 'T', 188-196, 'DKG', 200, 292-314, 'W', 316-318, 'X', 4396-4401 <EAT>
R;Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meier, K.; Schwartz, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A;Title: 5' control regions of the apolipoprotein(a) gene and members of the related plasma protein family. A;Reference number: A47277; MUID:93165698; PMID:7679504
A;Accession: A47277
A;Status: preliminary; translation not shown; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-16 <RES>
A;Cross-references: GB:I07899; NID:g967973; PID:g967974
R;Malgaretti, N.; Acquati, F.; Magneschi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Saccoccio, L.
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A;Title: Characterization by yeast artificial chromosome cloning of the linked apolipoprotein(a) gene and members of the related plasma protein family. A;Reference number: A47233; MUID:93087573; PMID:1454851
A;Accession: I60906
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE2>
A;Cross-references: GB:M90078; NID:g178786; PIDN:AAA35547.1; PID:g553188
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: A47233
A;Status: preliminary; translation not shown; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE5>
A;Cross-references: GB:M90079; NID:g178784; PIDN:AAA35546.1; PID:g553187
R;Ichinose, A.
Biochemistry 31, 3113-3118, 1992
A;Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated with atherosclerosis. A;Reference number: I52415; MUID:92207924; PMID:1554658
A;Accession: I52415
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-16 <RE3>
A;Cross-references: GB:M86877; NID:g178780; PIDN:AAB49909.1; PID:g553185
A;Note: apo(a) gene 1 (nomenclature of reference I52415)
A;Accession: I65286
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-16 <RB4>
A;Cross-references: GB:M86878; NID:g178782; PIDN:AAA51749.1; PID:g553186
C;Genetics:
A;Gene: GDB:LPA
A;Cross-references: GDB:I120699; OMIM:152200
A;Map position: 6q26-6q27
A;Note: several genes closely linked on chromosome 6 are identical in the first coding exons of kringle repeats
C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F;28-105/Domain: kringle homology <KR1>
F;142-219/Domain: kringle homology <KR2>
F;256-333/Domain: kringle homology <KR3>
F;370-447/Domain: kringle homology <KR4>
F;484-561/Domain: kringle homology <KR5>
F;598-675/Domain: kringle homology <KR6>
F;712-789/Domain: kringle homology <KR7>
F;826-903/Domain: kringle homology <KR8>
F;940-1017/Domain: kringle homology <KR9>
F;1054-1131/Domain: kringle homology <KR10>
F;1168-1245/Domain: kringle homology <KR11>
F;1282-1359/Domain: kringle homology <KR12>
F;1396-1473/Domain: kringle homology <KR13>
F;1510-1587/Domain: kringle homology <KR14>
F;1624-1701/Domain: kringle homology <KR15>
F;1738-1815/Domain: kringle homology <KR16>
F;1852-1929/Domain: kringle homology <KR17>
F;1966-2043/Domain: kringle homology <KR18>

F;2080-2157/Domain: kringle homology <KR19>
F;2194-2271/Domain: kringle homology <KR20>
F;2308-2385/Domain: kringle homology <KR21>
F;2422-2499/Domain: kringle homology <KR22>
F;2536-2613/Domain: kringle homology <KR23>
F;2650-2727/Domain: kringle homology <KR24>
F;2764-2841/Domain: kringle homology <KR25>
F;2878-2955/Domain: kringle homology <KR26>
F;2992-3069/Domain: kringle homology <KR27>
F;3106-3183/Domain: kringle homology <KR28>
F;3220-3297/Domain: kringle homology <KR29>
F;3334-3411/Domain: kringle homology <KR30>
F;3448-3525/Domain: kringle homology <KR31>
F;3562-3639/Domain: kringle homology <KR32>
F;3676-3753/Domain: kringle homology <KR33>
F;3782-3859/Domain: kringle homology <KR34>
F;3896-3973/Domain: kringle homology <KR35>
F;4010-4087/Domain: kringle homology <KR36>
F;4124-4201/Domain: kringle homology <KR37>
F;4228-4307/Domain: kringle homology <KR38>
F;4328-4541/Domain: trypsin homology <TRY>

Query Match 83.9%; Score 933.5; DB 1; Length 4548;
Best Local Similarity 84.5%; Pred. No. 6.4e-78;
Matches 174; Conservative 11; Mismatches 12; Indels 9; Gaps 1;

Qy 2 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFETPRK 61
Db 4352 HFCGGTLLSPWVLTAAHCKLKSSRPSSYKVLGAHQEVNLESHVQIEVSRFLFETQA 4411

Qy 62 DIALLKLSPPAVITDKVIPACLPSPNYVADRTCEFTTGWGETGTGAGLLKEAQLPVI 121
Db 4412 DIALLKLSPPAVITDKVMPACLPSPDYVMTARTCYITGWGETGTGAGLLKEAQLVI 4471

Qy 122 ENKVCNRYEFLNGRVQSTELCAGHAGTSDCGDGGPLVCFKDKYILQGVTSWGLGC 181
Db 4472 ENEVCNRYK-----TCAEHLARGTSDCGDGGPLVCFKDKYILQGVTSWGLGC 4522

Qy 182 ARPNKPGVYVRVSRFVTWIEGVMRNN 207
Db 4523 ARPNKPGVYVRVSRFVTWIEGVMRNN 4548

RESULT 6
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N;Contains: angiotensin; plasminogen
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: A38514; S48202; S48203
R;Degeen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A;Title: Characterization of the cDNA coding for mouse plasminogen and localization of b
A;Reference number: A38514; MUID:91184812; PMID:2081600
A;Accession: A38514
A;Molecule type: mRNA
A;Residues: 1-812 <DBG>
A;Cross-references: GB:J04766; NID:G200402; PIDN:AAA50168.1; PID:G200403
R;Liijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076; PMID:7523120
A;Accession: S48202
A;Molecule type: protein
A;Residues: 20-25 <LIJ>
A;Accession: S48203
A;Molecule type: protein
A;Residues: 22-27 <LIJ>
C;Comment: plasminogen is synthesized by the kidney and is present in plasma and many ot
mediately after dissociation from the clot. In the presence of the inhibitor, the activa
e inhibitor, the activation involves also removal of the activation peptide.
C;Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiotensin. To

eful in treating solid tumors.
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydr
F;1-96/Domain: plasminogen-related protein precursor homology <SIG>
F;1-19/Domain: signal sequence #status predicted <PRO>
F;20-812/Product: plasminogen #status predicted <PRO>
F;20-96/Domain: activation peptide #status predicted <APT>
F;79-466/Product: angiotensin #status predicted <AST>
F;97-581,582-812/Product: plasmin #status predicted <MAT>
F;97-581/Domain: chain A #status predicted <ACH>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;582-812/Domain: chain B #status predicted <BCH>
F;582-805/Domain: trypsin homology <TRY>
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32
bonds: #status predicted
F;78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F;136,308/Binding site: carbonylhydride (Asn) (covalent) #status predicted
F;466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F;581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 83.1%; Score 925; DB 1; Length 812;
Best Local Similarity 84.0%; Pred. No. 4.9e-78;
Matches 173; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFETPRK 61
Db 607 HFCGGTLLSPWVLTAAHCKLEKSRPEFYKVLGAHEEYIRGLDVQISVAKLLLEPNR 666

Qy 62 DIALLKLSPPAVITDKVIPACLPSPNYVADRTCEFTTGWGETGTGAGLLKEAQLPVI 121
Db 667 DIALLKLSPPAVITDKVIPACLPSPNYVADRTCYITGWGETGTGAGLLKEAQLPVI 726

Qy 122 ENKVCNRYEFLNGRVQSTELCAGHAGTSDCGDGGPLVCFKDKYILQGVTSWGLGC 181
Db 727 ENKVCNRYEFLNGRVQSTELCAGHAGTSDCGDGGPLVCFKDKYILQGVTSWGLGC 786

Qy 182 ARPNKPGVYVRVSRFVTWIEGVMRNN 207
Db 787 ARPNKPGVYVRVSRFVTWIEGVMRNN 812

RESULT 7
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C;Accession: B61545; S28200
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015; PMID:3168975
A;Accession: B61545
A;Molecule type: protein
A;Residues: 1-37,38-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A;Title: Complete amino acid sequence of ovine miniplasminogen.
A;Reference number: S28200; MUID:93149995; PMID:1492092
A;Accession: S28200
A;Molecule type: protein
A;Residues: 118-460 <SC>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology

C;keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase;
 F;1-37,38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>
 F;1-37/Domain: activation peptide (fragment) #status experimental <APT>
 F;38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
 F;41-118/Domain: kringle homology <KR4>
 F;118-460/Product: miniplasminogen #status experimental <MIN>
 F;132-211/Domain: kringle homology <KR5>
 F;226-460/Domain: plasmin chain B #status experimental <BCH>
 F;231-453/Domain: trypsin homology <TRY>
 F;272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 82.9%; Score 923; DB 2; Length 460;
 Best Local Similarity 84.8%; Pred. No. 3.8e-78;
 Matches 173; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPDWLTAACHLEKSPSPSSYKVLGAHQEVNLEPHVQIEVSRLFLPEPTRK 61
 Db 255 HFCGGTLLSPDWLTAACHLEKSPSPSSYKVLGAHQEVNLEPHVQIEVSRLFLPEPTRK 314
 Qy 62 DIALLKLSPPAVITDKVIPACLPSPNNVADRTCEFTTGMGTGAGLLKEAQLPVI 121
 Db 315 DIALLKLSPPAVITDKVIPACLPSPNNVADRTCEFTTGMGTGAGLLKEAQLPVI 374
 Qy 122 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKKYILQGVTSWGLGC 181
 Db 375 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKKYILQGVTSWGLGC 434
 Qy 182 ARPNKPGVYVRVSRFTWIEGVNRN 205
 Db 435 ARPNKPGVYVRVSRFTWIEGVNRN 458

RESULT 8

plasmin (EC 3.4.21.7) precursor - horse (fragments)
 N;Alternate names: miniplasminogen
 C;Species: Equus caballus (domestic horse)
 C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
 C;Accession: A61545; S17527
 R;Schaller, J.; Rickli, E.E.
 Enzyme 40, 63-69, 1988
 A;Title: Structural aspects of the plasminogen of various species.
 A;Reference number: A61545; MUID:99005015; PMID:3168975
 A;Accession: A61545
 A;Molecule type: protein
 A;Residues: 1-33;34-117 <SCH>
 R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
 Protein Seq. Data Anal. 4, 69-74, 1991
 A;Title: Complete amino acid sequence of equine miniplasminogen.
 A;Reference number: S17527; MUID:92052077; PMID:1946332
 A;Accession: S17527
 A;Molecule type: protein
 A;Residues: 118-455 <SC2>
 C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase;
 F;1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
 F;1-33/Domain: activation peptide (fragment) #status experimental <APT>
 F;34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MAT>
 F;37-114/Domain: kringle homology <KR4>
 F;118-455/Product: miniplasminogen #status experimental <MIN>
 F;126-205/Domain: kringle homology <KR5>
 F;226-455/Domain: plasmin chain B #status experimental <BCH>
 F;226-448/Domain: trypsin homology <TRY>
 F;267,310,405/Active site: His, Asp, Ser #status predicted

Query Match 81.3%; Score 905; DB 2; Length 455;
 Best Local Similarity 79.1%; Pred. No. 1.8e-76;
 Matches 163; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPDWLTAACHLEKSPSPSSYKVLGAHQEVNLEPHVQIEVSRLFLPEPTRK 61
 Db 250 HFCGGTLLSPDWLTAACHLEKSPSPSSYKVLGAHQEVNLEPHVQIEVSRLFLPEPTRK 309

Qy 62 DIALLKLSPPAVITDKVIPACLPSPNNVADRTCEFTTGMGTGAGLLKEAQLPVI 121
 Db 310 DIALLKLSPPAVITDKVIPACLPSPNNVADRTCEFTTGMGTGAGLLKEAQLPVI 369
 Qy 122 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKKYILQGVTSWGLGC 181
 Db 370 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKKYILQGVTSWGLGC 429
 Qy 182 ARPNKPGVYVRVSRFTWIEGVNRN 207
 Db 430 ARPNKPGVYVRVSRFTWIEGVNRN 455

RESULT 9

I46260
 plasmin (EC 3.4.21.7) precursor - western European hedgehog
 C;Species: Erinaceus europaeus (western European hedgehog)
 C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
 C;Accession: I46260
 R;Jawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
 J. Biol. Chem. 270, 24004-24009, 1995
 A;Title: The recurring evolution of Ip(a): Insights from cloning of hedgehog apolipoprotein
 A;Reference number: I46259; MUID:96025778; PMID:7592597
 A;Accession: I46260
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-810 <LAW>
 A;Cross-references: EMBL:U33171; NID:g1046360; PID:g1046361
 C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology,
 C;Keywords: hydrolase; serine proteinase
 F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
 F;103-181/Domain: kringle homology <KR1>
 F;185-262/Domain: kringle homology <KR2>
 F;275-352/Domain: kringle homology <KR3>
 F;379-456/Domain: kringle homology <KR4>
 F;482-561/Domain: kringle homology <KR5>
 F;582-803/Domain: trypsin homology <TRY>

Query Match 80.1%; Score 891; DB 2; Length 810;
 Best Local Similarity 79.1%; Pred. No. 7.2e-75;
 Matches 163; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPDWLTAACHLEKSPSPSSYKVLGAHQEVNLEPHVQIEVSRLFLPEPTRK 61
 Db 605 HFCGGTLLSPDWLTAACHLEKSPSPSSYKVLGAHQEVNLEPHVQIEVSRLFLPEPTRK 664
 Qy 62 DIALLKLSPPAVITDKVIPACLPSPNNVADRTCEFTTGMGTGAGLLKEAQLPVI 121
 Db 665 DIALLKLSPPAVITDKVIPACLPSPNNVADRTCEFTTGMGTGAGLLKEAQLPVI 724
 Qy 122 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKKYILQGVTSWGLGC 181
 Db 725 ENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKKYILQGVTSWGLGC 784
 Qy 182 ARPNKPGVYVRVSRFTWIEGVNRN 207
 Db 785 ARPNKPGVYVRVSRFTWIEGVNRN 810

RESULT 10

PLBO
 plasmin (EC 3.4.21.7) precursor - bovine
 N;Alternate names: plasminogen
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
 C;Accession: S45046; A25835; I45961; S03736
 R;Berglund, L.; Andersen, M.D.; Petersen, T.E.
 submitted to the EMBL Data Library, May 1994
 A;Description: Cloning and characterization of the bovine plasminogen cDNA.
 A;Reference number: S45046
 A;Accession: S45046
 A;Molecule type: mRNA

A;Residues: 1-812

A;Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963
A;Experimental source: liver
A;Note: it is uncertain whether Met-1 or Met-8 is the initiator
R;Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Roselet, S.J.; Kampfer, U.; Rick
Eur. J. Biochem. 149, 267-278, 1985
A;Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plas
A;Reference number: A25835; MUID:85203906; PMID:3846532
A;Accession: A25835
A;Molecule type: protein
A;Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A;Reference number: 145961; MUID:85023311; PMID:6148961
A;Accession: 145961
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 706-743, 'R', 745-812 <MAL>
A;Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552
R;Brundish, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A;Reference number: 803735; MUID:81212097; PMID:7238497
A;Accession: 803735
A;Molecule type: protein
A;Residues: 27-83 <BRU>
A;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
na the walls of the graafian follicle; also activates the urokinase-type plasminogen act
C;Pathway: fibrinolysis
C;Superfamily: plasmin;
C;Keywords: duplication; fibrinolysis; glycoprotein; hydrolyase; kidney; kringie; plasma;
F;1-26/Domain: signal sequence #status predicted <SIG>
F;8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F;27-812/Product: plasminogen #status experimental <PRO>
F;27-103/Domain: activation peptide #status experimental <APT>
F;104-583, 584-812/Product: chain A #status experimental <MAT>
F;104-583/Domain: plasmin chain A #status experimental <MAT>
F;110-188/Domain: kringie homology <KR1>
F;192-269/Domain: kringie homology <KR2>
F;282-359/Domain: kringie homology <KR3>
F;384-461/Domain: kringie homology <KR4>
F;485-564/Domain: kringie homology <KR5>
F;584-812/Domain: plasmin chain B #status experimental <BCH>
F;584-805/Domain: trypsin homology <TRY>
F;56-80, 60-68, 110-186, 131-171, 159-183, 192-269, 195-323, 213-252, 241-264, 282-359, 303-342, 33
bonds: #status predicted
F;315/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;624, 667, 762/Active site: His, Asp, Ser #status predicted

Query Match 78.9%; Score 878; DB 1; Length 812;
Best Local Similarity 80.1%; Pred. No. 1.2e-73;
Matches 165; Conservative 12; Mismatches 29; Indels 0; Gaps 0;
Qy 2 HFCGGTILSPFWLTAACHLSEKSPSSYKVLGAHQVNLEPHVQIEVSRLEPRK 61
Db 607 HFCGGTILSPFWLTAACHLSEKSPSSYKVLGAHQVNLEPHVQIEVSRLEPRK 666
Qy 62 DIALLKLSPPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKBAQLPVI 121
Db 667 DIALLKLSRAIITKEVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKBAQLPVI 726
Qy 122 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCGDSGGPLVCFKDKYILQGVTSWGLGC 181
Db 727 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCGDSGGPLVCFKDKYILQGVTSWGLGC 786
Qy 182 ARPNKPGYVVRVSRFVTWIEGWMRN 207
Db 787 ARPNKPGYVVRVSRFVTWIEGWMRN 812

RESULT 11
S00845
hepsin (EC 3.4.21.-) - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 18-Jun-1999
C;Accession: S00845
R;Leytus, S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.
Biochemistry 27, 1067-1074, 1988
A;Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane dom
A;Reference number: S00845; MUID:88209431; PMID:2835076
A;Accession: S00845
A;Molecule type: mRNA
A;Residues: 1-417 <LSY>
A;Cross-references: EMBL:X07732; NID:G32063; PIDN:CAA30558.1; PID:G32064
C;Genetics:
A;Gene: GDB:HPN; TMPRSS1; hepsin
A;Cross-references: GDB:I35685; OMIM:142440
A;Map position: 19q11-19q13.2
C;Superfamily: hepsin; trypsin homology
C;Keywords: hydrolase; liver; serine proteinase; transmembrane protein
F;23-45/Domain: transmembrane #status predicted <TN>
F;163-400/Domain: trypsin homology <TRY>
F;188-204, 291-359, 322-338, 349-381/Disulfide bonds: #status predicted
F;203, 257, 353/Active site: His, Asp, Ser #status predicted

Query Match 39.9%; Score 444; DB 1; Length 417;
Best Local Similarity 42.2%; Pred. No. 1.5e-33;
Matches 92; Conservative 38; Mismatches 66; Indels 22; Gaps 7;
Qy 2 HFCGGTILSPFWLTAACHLSEKSPSSYKVLGAHQVNLEPHVQIEVSRLEPRK 54
Db 186 HLCGGSLSGDWLTAACHLSEKSPSSYKVLGAHQVNLEPHVQIEVSRLEPRK 243
Qy 55 --PLEPRK---DIALKLSPPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGT 108
Db 244 LPFRDPNSENNDIALVHLSPLPTEYTPVCLPAGQALVDGKICTVTGWNQY-Y 302
Qy 109 G--AGLLKBAQLPVIENTKCNRYEFLNGRVQSTELCAGHLAGTDSQCGDSGGPLVCFEK 166
Db 303 GQAGVLQEARVPIISNDVCGADFYGNQIKPKMFCAGYPEGIDACQDGSQGPFCRDS 362
Qy 167 ----DKYILQGVTSWGLGCARPNKPGYVVRVSRFVTW 200
Db 363 ISRTPRWRLCGIVSWGTCALAQKPGYVYTKVDFREWI 400

RESULT 12
S33777
hepsin (EC 3.4.21.-) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C;Accession: S33777; S32013
R;Farley, D.; Raymond, F.; Nick, H.
Biochim. Biophys. Acta 1173, 350-352, 1993
A;Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase.
A;Reference number: S33777; MUID:93305733; PMID:8318546
A;Accession: S33777
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-416 <FAR>
A;Cross-references: EMBL:X70900; NID:G57928; PIDN:CAA50256.1; PID:G57929
C;Superfamily: hepsin; trypsin homology
C;Keywords: hydrolase; liver; serine proteinase; transmembrane protein
F;22-44/Domain: transmembrane #status predicted <TN>
F;162-399/Domain: trypsin homology <TRY>
F;187-203, 290-358, 321-337, 348-380/Disulfide bonds: #status predicted
F;202, 256, 352/Active site: His, Asp, Ser #status predicted

Query Match 38.0%; Score 423; DB 1; Length 416;
Best Local Similarity 40.8%; Pred. No. 1.3e-31;
Matches 89; Conservative 37; Mismatches 70; Indels 22; Gaps 7;
Qy 2 HFCGGTILSPFWLTAACHLSEKSPSSYKVLGAHQVNLEPHVQIEVSRLEPRK 54

```

Db      185 HLCGSLSGDWLTAARCPERRNVLRSRVRPAGA--VARTSPHAVQLGQVAVIHYGGY 242
Qy      55 --FLEPT-----RKDIALLKSSPAVITDKVIPACLPSPNYVADRTSCFTGWTGQTGF 108
Db      243 LPFRDPTIDENSNDIALVHLSSSLPLETYIQVCLPAAGQALVDGKVCVTTVGWNTQ--FY 301
Qy      109 G--AGLAKAQLPVIENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQSGSGGPLVCFEX 166
Db      302 GQARVILQEARVPIISNEVCNRPDYGNQIKPKMFCAGYPEGGIDACQSGSGGHFVCEDR 361
Qy      167 ----DKTILQGVTSWGLGCLARPKNPGVYVRSRVFTWI 200
Db      362 ISGTSRRLCGIVSMGTGCLARKEPGVYTKVIDPREMI 399

RESULT 13
A43090
enteropeptidase (EC 3.4.21.9) precursor [validated] - bovine
N;Alternate names: enterokinase
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Apr-2003
C;Accession: A43090 #sequence_revision 10-Sep-1999 #text_change 28-Apr-2003
R;Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
A;Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of two distinct regions
A;Reference number: A43090; MUID:94322561; PMID:8052624
A;Accession: A43090
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1035 <KIT>
A;Cross-references: GB:U09859; NID:G746410; PIDN:AAB40026.1; PID:G746411
A;Experimental source: small intestine
R;Lavallie, E.R.; Rehmetulla, A.; Racte, L.A.; DiBlasio, E.A.; Ferenz, C.; Grant, K.L.; J. Biol. Chem. 268, 23311-23317, 1993
A;Title: Cloning and functional expression of a cDNA encoding the catalytic subunit of human enterokinase
A;Reference number: A48874; MUID:94043122; PMID:8226855
A;Accession: A48874
A;Molecule type: mRNA
A;Residues: 801-1035 <LAV>
A;Cross-references: GB:L19663; NID:G416131; PIDN:AAA16035.1; PID:G416132
A;Note: parts of this sequence, including the amino end of the mature protein, were confirmed by N-terminal sequencing
R;Light, A.; Janska, H.
J. Protein Chem. 10, 475-480, 1991
A;Title: The amino-terminal sequence of the catalytic subunit of bovine enterokinase.
A;Reference number: A61436; MUID:92189715; PMID:1799406
A;Accession: A61436
A;Molecule type: protein
A;Residues: 801-807, 'y', 809-827 <LIG>
C;Comment: The mechanism of association with the membrane of the intestinal brush border enzyme attachment using a signal-anchor sequence
C;Comment: Conversion from membrane-bound to soluble forms may involve further processing
C;Complex: mature enteropeptidase is variously reported to contain two (heavy and light) subunits
C;Function:
A;Description: cleaves propeptide from trypsinogen to produce active trypsin
A;Pathway: intestinal digestive hydrolase cascade
C;Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding repeat homology
F;22-38/Domain: transmembrane #status predicted <TM>
F;52-117/Product: enteropeptidase mini chain #status predicted <MCH>
F;118-800/Product: enteropeptidase heavy chain #status predicted <HCH>
F;199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;358-520/Domain: MAM homology <MAM>
F;542-647/Domain: C1r/C1s repeat homology <C1R>
F;659-693/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;694-799/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC>
F;801-1035/Product: enteropeptidase light chain #status predicted <LCH>
F;801-1030/Domain: trypsin homology <TRY>
F;116,147,170,194,233,263,284,404,456,486,519,550,646,698,722,741,762,864,903,965/Binding site: active site
F;788-912,826-842,926-993,957-972,983-1011/Disulfide bonds: #status predicted
F;841,892,987/Active site: His, Asp, Ser #status predicted

```

```

Query Match      38.0%; Score 423; DB 1; Length 1035;
Best Local Similarity 39.8%; Pred. No. 4e-31;
Matches 84; Conservative 41; Mismatches 74; Indels 12; Gaps 6;

Qy      4 CGGTLSPEWVLTAAHCL-EKSPRPSSYKYLGAHQEVNL-EPHVQIEVSRLEP--- 58
Db      826 CGASLSVRDLVLSAAHCVYGRNMEPSKKAIVLGLHMASNLTSPOIETRLIDQIVINPHYN 885

Qy      59 ---TRDIALLKSSPAVITDKVIPACLPSPNYVADRTSCFTGWTGSET--QGTFFGAGLL 113
Db      886 KRRNNDIAMHLEMKVNYTDYIQICLPENQVFPFPRICISAGWAGALYQGS-TADVL 944

Qy      114 KEAQLPVIENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQSGSGGPLVCFEKDKYILQG 173
Db      945 QEADVLLSNEKCOQ-QMPEYNITENNVCAGYEAGGVDSQSGSGGPLMCOENNRWLLAG 1003

Qy      174 VTSGLGCLARPKNPGVYVRSRVFTWIEGMV 204
Db      1004 VTSFGYQCALPNRPFGVYVRSRVFTWIEQSPL 1034

RESULT 14
A47246
tryptase (EC 3.4.21.59) 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A47246
R;McNeill, H.P.; Reynolds, D.S.; Schiller, V.; Ghildyal, N.; Gurley, D.S.; Austen, K.F.; et al.
Proc. Natl. Acad. Sci. U.S.A. 89, 11174-11178, 1992
A;Title: Isolation, characterization, and transcription of the gene encoding mouse mast cell tryptase
A;Reference number: A47246; MUID:93087489; PMID:1454796
A;Accession: A47246
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-273 <MCN>
A;Cross-references: GB:L00653; NID:G200518; PIDN:AAA39992.1; PID:G200519
A;Note: sequence extracted from NCBI backbone (NCBIN:119745, NCBIP:119746)
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;23-285/Domain: trypsin homology <TRY>

Query Match      37.6%; Score 419; DB 2; Length 273;
Best Local Similarity 43.4%; Pred. No. 1.9e-31;
Matches 95; Conservative 33; Mismatches 65; Indels 26; Gaps 9;

Qy      1 MHFCGGTLLSPWVLTAAHCLKS-PPSPSSYKYLGAHQEVNL-EPHVQIEVSRLEP--- 58
Db      54 MHFCGGSLIHPQWVLTAAHCVGPDVADPNKRVQL---RKQYLYYHDLMTVSGIITHPD 110

Qy      59 ----TRKDIALLKSSPAVITDKVIPACLPSPNYVADRTSCFTGWTGSETGTFGAGL- 112
Db      111 FYIVQDGADIALKLINPNVNSIDYVHPVLPFPASETFPSGTLCWVWGNIIDN--GVNLP 168

Qy      113 ----LKEAQLPVIENKVCN-RYE--FLNGR-----VOSTELCAGHLAGGTDSCQSGSGGPL 161
Db      169 PPFPLEKEVQVPIIENHLCDLKXHKGLITGDNVHVRDMLCAGN--EGHDSQCGDSGGPL 226

Qy      162 VCEKDKYILQGVTSWGLGCLARPKNPGVYVRSRVFTWI 200
Db      227 VCKVEDTFLQAGVSVSGEGCAQNPFRGIYTRVYYLDWI 265

RESULT 15
A56318
enteropeptidase (EC 3.4.21.9) precursor [validated] - human
N;Alternate names: enterokinase
C;Species: Homo sapiens (man)
C;Date: 19-May-1995 #sequence_revision 09-Aug-1996 #text_change 28-Apr-2003
C;Accession: A56318; B43090
R;Kitamoto, Y.; Velte, R.A.; Donis-Keller, H.; Sadler, J.E.
Biochemistry 34, 4562-4568, 1995
A;Title: cDNA sequence and chromosomal localization of human enterokinase, the proteolytic activator of trypsinogen
A;Reference number: A56318; MUID:95234679; PMID:7718557

```


A:Accession: A56318
A:Molecule type: mRNA
A:Residues: 1-1019 <KIT>
A:Cross-references: GB:U09860; NID:g746412; PIDN:AAC50138.1; PID:g746413
R:Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
A:Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of two distinct domains
A:Reference number: A43090; MUID:94329561; PMID:8052624
A:Accession: B43090
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 749-1019 <KI2>
A:Cross-references: GB:U09860
C:Comment: The mechanism of association with the membrane of the intestinal brush border is not clear. The enzyme is a mosaic protease composed of two distinct domains (indicated below) or with amino-terminal myristoylation of the heavy chain.
C:Genetics:
A:Gene: GDB:PRSS7
A:Cross-references: GDB:384083; OMIM:226200
A:Map position: 21q21-21q21
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and light) subunits by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involves cleavage of the heavy chain.
C:Function:
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
A:Pathway: intestinal digestive hydrolase cascade
C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding repeat
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
F:1-784/Product: enteropeptidase heavy chain #status predicted <HCH>
F:22-38/Domain: transmembrane #status predicted <TM>
F:184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:342-504/Domain: MAM homology <MAM>
F:526-631/Domain: C1r/C1s repeat homology <C1R>
F:643-773/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:678-783/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC>
F:785-1019/Product: enteropeptidase light chain #status predicted <LCH>
F:785-1014/Domain: trypsin homology <TRY>
F:116,147,179,328,388,440,470,503,534,630,682,706,725,848,887,909,949/Binding site:
F:772-896,810-826,910-977,941-956,967-995/Disulfide bonds: #status predicted
F:825,876,971/Active site: His, Asp, Ser #status predicted

Query Match 37.6%; Score 418; DB 1; Length 1019;
Best Local Similarity 41.2%; Pred. No. 1.1e-30;
Matches 87; Conservative 36; Mismatches 76; Indels 12; Gaps 7;

Qy 4 CGGTLISPEWVLTAAHCL-EKSPRSSYKVLGAHQEVLN-EPHVQIEIVSRLEFLEP--- 58
Db 810 CGASIVSDMLVSAHCVYGRNLEPSKWTAILGLMKSNLTSPQTVPRLIDEIVINPHYN 869

Qy 59 -TRK-DIALKLSPAVITDKVIPACLPSPNYVADRTECFITGWGET--QGTFGAGLL 113
Db 870 RRRKNDIAMHLEFKVNYTDYIQICLPEENQVPPGRCNSIAGWGTVVYQGT-TANIL 928

Qy 114 KEAQLPVIEKVCNRYEFLNGRVOSTELCAGHLAGGTTSCQDSGGPLVCFEKDKYIIQG 173
Db 929 QEADVPLLSNRCQQ-QMPEYNITENNICAGYEGGIDSCQDSGGPLMCOENNRWFLAG 987

Qy 174 VTSWGLGCARPKGVYVVRVFRVFTWIEGVN 204
Db 988 VTSFGYKALPNRFGVAVRSRFTETWISQFL 1018

Search completed: September 10, 2004, 14:55:47

Job time : 41 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 14:38:48 ; Search time 127 Seconds
(without alignments)
460.530 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113

Sequence: 1 MHFCGGTLISPEWVLTAAHC.....GVYVRVSRFTWIEGVNRNN 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1113	100.0	207	6	ABR48479 Human pla
2	1113	100.0	230	2	RAY31159 Human pla
3	1113	100.0	249	5	ABB81497 Human mic
4	1113	100.0	348	5	ABB81498 Human min
5	1113	100.0	437	2	AAW51457 Human pla
6	1113	100.0	546	2	AAR34427 Sequence
7	1113	100.0	790	2	AAR60519 Human 'G1
8	1113	100.0	790	4	AAB36562 Mammalian
9	1113	100.0	791	3	RAY50867 Human pla
10	1113	100.0	791	3	AAO1887 Human pla
11	1113	100.0	791	3	RAY9589 Human pla
12	1113	100.0	791	4	AKG67223 Amino aci
13	1113	100.0	791	5	AAE23660 Human pla
14	1113	100.0	791	6	ABG76087 Human pla
15	1113	100.0	791	6	ADA08448 Mammalian
16	1113	100.0	810	2	AAR08065 Human pla
17	1113	100.0	810	2	AAR13221 Human pla
18	1113	100.0	810	2	AAR13219 R561G hum
19	1113	100.0	810	2	AAR12406 R561E hum
20	1113	100.0	810	2	AAR13220 R561S hum
21	1113	100.0	810	2	AAR34428 Sequence
22	1113	100.0	810	2	AAW31169 Plasmidog
23	1113	100.0	810	2	RAY02114 SEQ ID 77
24	1113	100.0	810	2	AY08685 Human pla
25	1113	100.0	810	3	AY82690 Human pla

26	1113	100.0	810	3	RAY53867 Amino aci
27	1113	100.0	810	5	ABB81496 Human pla
28	1113	100.0	810	5	ABB83795 Human pla
29	1113	100.0	810	6	ABG75602 Human pla
30	1113	100.0	810	6	AAE37127 Human pla
31	1113	100.0	810	6	AAE36399 Human pla
32	1112	99.9	810	5	ABB83470 Human pla
33	1112	99.9	810	5	ABB75939 Human pla
34	1112	99.9	1220	4	AAU32858 Novel hum
35	1108	99.6	791	2	AAW34285 Human pla
36	1106	99.4	230	2	AAR56474 Mutant pl
37	1106	99.4	230	2	AAR49031 Plasmid p
38	1106	99.4	230	2	AAR56473 Mutant pl
39	1106	99.4	230	2	AAR56472 Plasmid p
40	1106	99.4	243	2	AAR54766 Serine pr
41	1106	99.4	810	2	AAR12938 Plasmidog
42	1106	99.4	810	5	ABB75947 Human pla
43	1106	99.4	811	2	AAR12943 Plasmidog
44	1106	99.4	811	2	AAR12945 Plasmidog
45	1106	99.4	811	2	AAR12944 Plasmidog

ALIGNMENTS

RESULT 1
ABR48479
ID ABR48479 standard; protein; 207 AA.
XX
AC ABR48479;
XX
DT 13-JUN-2003 (first entry)
XX
DE Human Plasmidote.
XX
KW Human; GENSET; therapeutic; therapy.
XX
OS Homo sapiens.
XX
PN WO200294864-A2.
XX
PD 28-NOV-2002.
XX
PF 06-AUG-2001; 2001WO-IB001715.
XX
PR 25-MAY-2001; 2001US-0293574P.
PR 15-JUN-2001; 2001US-0298698P.
PR 29-JUN-2001; 2001US-0302277P.
PR 13-JUL-2001; 2001US-0305456P.
XX
(GEBT) GENSET.
PI Bejanin S, Tanaka H;
XX
DR WPI; 2003-129412/12.
DR N-PSDB; ACCS1086.
XX
PT New GENSET polynucleotides and polypeptides, useful for preparing a
composition for treating GENSET-related disorders and as reagents in
assays to quantitatively determined levels of GENSET expression in
biological samples.
XX
PS Claim 2; Page 447-448; 505pp; English.
XX
CC The present invention relates to novel human GENSET coding sequences
(ACCS1060-ACCS1115) and proteins (ABR48453-ABR48508). The GENSET
sequences are useful for preparing a composition for treating GENSET-
related disorders. They can also be used as markers for tissues in which
the corresponding protein is preferentially expressed, as molecular
weight markers on Southern gels, as chromosome markers or tags to
identify chromosomes, and as reagents in assays to quantitatively
determined levels of GENSET expression in biological samples
XX

```
SQ Sequence 207 AA;
Query Match 100.0%; Score 1113; DB 6; Length 207;
Best Local Similarity 100.0%; Pred. No. 9.7e-108; Mismatches 0; Indels 0; Gaps 0;
Matches 207; Conservative 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSVKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 1 MHFCGGTLLSPWVLTAAHCLKSPSSVKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 120
Db 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 120
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 180
Db 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 180
Qy 181 CARPNKPGVYVRSRVFTWIEGVNRN 207
Db 181 CARPNKPGVYVRSRVFTWIEGVNRN 207

RESULT 2
AAY31159
ID AAY31159 standard; protein; 230 AA.
XX
AC AAY31159;
XX
DT 26-OCT-1999 (first entry)
XX
DE Human plasminogen serine protease protein domain.
XX
KW Macrophage stimulating protein; MSP; human; modulator; proliferation;
KW differentiation; intestinal epithelium; colon crypt; treatment; cancer;
KW haematopoietic disorder; megakaryocyte deficiency; gastrointestinal;
KW chemotherapeutic agent; gut toxicity; serine protease; plasminogen.
XX
OS Homo sapiens.
XX
PN US5948892-A.
XX
PD 07-SEP-1999.
XX
PF 16-DEC-1996; 96US-00766982.
XX
PR 16-DEC-1996; 96US-00766982.
XX
PA (AMGE-) AMGEN INC.
XX
PI Wahl RC;
XX
WPI; 1999-517975/43.
XX
Analogues of macrophage stimulating protein for treating gastrointestinal
or haematopoietic disorders.
XX
Example 2; Col 27-28; 23pp; English.
XX
This invention describes a novel purified and isolated analogue of mature
macrophage stimulating protein (MSP) having at least one unpaired
cysteine residue substituted with another amino acid which modulates the
proliferation or differentiation of the intestinal epithelium. The
product of the invention binds to RON (a cell membrane protein tyrosine
kinase which is a member of the c-met family) to promote the formation of
colon crypts. MSP analogues are useful for the treatment of conditions
requiring the administration of MSP, such conditions include
haematopoietic disorders such as those involving a deficiency of
megakaryocytes and gastrointestinal disorders such as ulcerative colitis,
Crohn's disease and infections. The MSP analogues are useful for
maintaining and repairing the epithelial lining in the treatment of
cancer, where the aggressive use of chemotherapeutic agents or the use of
whole body radiation may lead to gut toxicity. The MSP analogues, which
```

```
CC have a higher activity than normal human MSP are effective at smaller
CC dosages, or optionally, they may be administered less frequently than
CC human MSP. This sequence represents a human plasminogen serine protease
CC domain which is used in a description of the method of the invention
XX
SQ Sequence 230 AA;
Query Match 100.0%; Score 1113; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.1e-107; Mismatches 0; Indels 0; Gaps 0;
Matches 207; Conservative 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSVKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 24 MHFCGGTLLSPWVLTAAHCLKSPSSVKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 83
Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 120
Db 84 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 143
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 180
Db 144 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 203
Qy 181 CARPNKPGVYVRSRVFTWIEGVNRN 207
Db 204 CARPNKPGVYVRSRVFTWIEGVNRN 230

RESULT 3
ABB81497
ID ABB81497 standard; protein; 249 AA.
XX
AC ABB81497;
XX
DT 02-SEP-2002 (first entry)
XX
DE Human microplasminogen protein SEQ ID NO:4.
XX
KW Human; plasminogen; microplasminogen; miniplasminogen; yeast; vector;
KW yeast expression vector; cardiant; thrombolytic; cerebroprotective;
KW fibrin proteolysis; kringle domain; thromboembolic disease;
KW focal cerebral ischaemic infarction; ischaemic stroke;
KW arterial thrombotic disease; peripheral arterial occlusive disease;
KW acute myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200250290-A1.
XX
PD 27-JUN-2002.
XX
PF 20-DEC-2001; 2001WO-BE000217.
XX
PR 21-DEC-2000; 2000GB-00031196.
XX
PR 09-JUL-2001; 2001GB-00016690.
XX
PR 09-JUL-2001; 2001GB-00016702.
XX
PA (THRO-) THROMB-X NV.
XX
PI Collen DJ, Nagai N, Laroche Y;
XX
WPI; 2002-500632/53.
XX
DR N-PSDB; ABN89460.
XX
Novel expression vector for expressing mammalian plasminogen derivatives
PT in yeast, has nucleotide sequence coding for catalytic domain of
PT plasminogen and/or coding for kringle domains of plasminogen linked to
PT promoter.
XX
Claim 30; Fig 3; 61pp; English.
XX
The present invention describes a yeast expression vector (I) comprising
CC a mammalian nucleotide sequence operably linked to a promoter, where the
```

CC mammalian nucleotide sequence codes for the catalytic domain of
CC plasminogen and further optionally codes for one or more kringle domains
CC of plasminogen, its mutants or hybrids. (I) has cardiant, thrombolytic
CC and cerebroprotective activities, and can mediate fibrin proteolysis.
CC (I) can be used useful for treating a thromboembolic disease in a mammal.
CC The mammalian protein expressed by (I) is useful for treating focal
CC cerebral ischaemic infarction (ischaemic stroke) or arterial thrombotic
CC diseases such as peripheral arterial occlusive disease or acute
CC myocardial infarction. The present sequence represents human
CC microplasminogen, which is used in the exemplification of the present
CC invention
CC
XX
SQ Sequence 249 AA;

RESULT 4	
ABB81498	
ID	ABB81498 standard; protein; 348 AA.
XX	
AC	ABB81498;
XX	
DT	02-SEP-2002 (first entry)
XX	
DE	Human miniplasminogen protein SEQ ID NO:6.
XX	
KW	Human; plasminogen; microplasminogen; miniplasminogen; yeast; vector;
KW	yeast expression vector; cardiant; thrombolytic; cerebroprotective;
KW	fibrin proteolysis; kringle domain; thromboembolic disease;
KW	focal cerebral ischaemic infarction; ischaemic stroke;
KW	arterial thrombotic disease; peripheral arterial occlusive disease;
KW	acute myocardial infarction.
XX	
OS	Homo sapiens.
XX	
PN	WO200250290-A1.
XX	
PD	27-JUN-2002.
XX	
PF	20-DEC-2001; 2001WO-BE000217.
XX	
PR	21-DEC-2000; 2000GB-00031196.
PR	09-JUL-2001; 2001GB-00016690.
PR	09-JUL-2001; 2001GB-00016702.
XX	
PA	(THRO-) THROMB-X NV.
XX	
PI	Collen DJ, Nagai N, Laroche Y;
XX	
DR	WPI; 2002-500632/53.
DR	N-PSDB; ABN89461.
XX	
PT	Novel expression vector for expressing mammalian plasminogen derivatives
PT	in yeast, has nucleotide sequence coding for catalytic domain of

plasminogen and/or coding for kringle domains of plasminogen linked to promoter.

Claim 31; Fig 4; 61pp; English.

The present invention describes a yeast expression vector (I) comprising a mammalian nucleotide sequence operably linked to a promoter, where the mammalian nucleotide sequence codes for the catalytic domain of plasminogen and further optionally codes for one or more kringle domains of plasminogen, its mutants or hybrids. (I) has cardiant, thrombolytic and cerebroprotective activities, and can mediate fibrin proteolysis. (I) can be used useful for treating a thromboembolic disease in a mammal. The mammalian protein expressed by (I) is useful for treating focal cerebral ischaemic infarction (ischaemic stroke) or arterial thrombotic diseases such as peripheral arterial occlusive disease or acute myocardial infarction. The present sequence represents human miniplasminogen, which is used in the exemplification of the present invention

RESULT 5
AAW51457
ID AAW51457 standard; protein; 437 AA.
XX
AC AAW51457;
XX
DT 02-SEP-1998 (first entry)
XX
DE Human plasminogen fragment with neovascularisation inhibiting activity.
XX
KW Human plasminogen; neovascularisation; angiostatin; inhibition; elastase; Sephacrose.
KW
OS Homo sapiens.
XX
Key Location/Qualifiers
FH Disulfide-bond 4..81
FT /label= Disulphide_bond
FT Disulfide-bond 25..64
FT /label= Disulphide_bond
FT Disulfide-bond 53..76
FT /label= Disulphide_bond
FT Disulfide-bond 108..187
FT /label= Disulphide_bond
FT Disulfide-bond 129..170
FT /label= Disulphide_bond
FT Disulfide-bond 158..182
FT /label= Disulphide_bond
FT Disulfide-bond 194..313
FT /label= Disulphide_bond
FT Disulfide-bond 204..212
FT /label= Disulphide_bond

OS Homo sapiens.
 XX WO9418227-A2.
 PN XX
 XX 18-AUG-1994.
 PD XX
 XX 04-FEB-1994; 94WO-DK000054.
 PF XX
 XX 04-FEB-1993; 93DK-00000130.
 PR XX
 XX 05-FEB-1993; 93DK-00000139.
 PR XX
 XX 03-DEC-1993; 93WO-GB002492.
 XX XX
 XX (DENZ-) DENZYME APS.
 PA XX
 XX Thogersen HC, Holtet TL, Etzerodt M;
 PI XX
 XX WPI; 1994-279681/34.
 DR XX
 XX Refolding of polypeptide molecules - using a cyclic process involving
 PT denaturing and renaturing conditions to produce a correctly folded prod.
 PT XX
 XX Disclosure; Page 148-50; 202pp; English.
 PS XX
 XX cDNA encoding kringle domains 1 and 4 of human plasminogen (full sequence
 CC given in AAR60519) was PCR amplified using primers given in AAQ71268-71.
 CC Amplified cDNA was linked to a sequence encoding the Factor-Xa cleavage
 CC site (given in AAR60503), subcloned in vector pLCIIMLCH6 so that it was
 CC linked to a hexahistidine-encoding sequence and expressed in E. coli
 CC QX13. The fusion protein was purified on an Ni2+-activated NTA-agarose
 CC column. A cyclic procedure was used to obtain correctly folded
 CC recombinant protein. (Updated on 25-MAR-2003 to correct PN field.)
 XX XX
 XX Sequence 790 AA;
 SQ
 Query Match 100.0%; Score 1113; DB 2; Length 790;
 Best Local Similarity 100.0%; Pred. No. 5.6e-107;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MHFCGGTLLSPFWLTAACHLEKSPRESSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60
 Db 584 MHFCGGTLLSPFWLTAACHLEKSPRESSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 643
 Qy 61 KDIALKLLSPAVITDKVIPACLPSPNVVADRTCEFTGTGQTFGAGLLKEAQLPV 120
 Db 644 KDIALKLLSPAVITDKVIPACLPSPNVVADRTCEFTGTGQTFGAGLLKEAQLPV 703
 Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
 Db 704 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 763
 Qy 181 CARPNKPGVYVRSRFTWIEGVNRN 207
 Db 764 CARPNKPGVYVRSRFTWIEGVNRN 790
 RESULT 8
 AAB36562
 ID AAB36562 standard; protein; 790 AA.
 XX
 AC AAB36562;
 XX
 XX 09-MAR-2001 (first entry)
 DT XX
 XX Mammalian kringle 5 protein SEQ ID NO:1.
 DE XX
 XX Kringle 5; anti-angiogenic; modified; blood protein; anti-inflammatory;
 KW vasotrophic; cytostatic; antithrombotic; antipsoriatic; antidiabetic;
 KW antiarteriosclerotic; osteoplastic; angiogenesis inhibitor; angiogenesis;
 KW inflammatory disorder; inflammation; chronic articular rheumatism;
 KW psoriasis; diabetic retinopathy; neovascular glaucoma; restenosis;
 KW capillary proliferation; atherosclerotic plaque; osteoporosis; cancer;
 KW solid tumour; angiofibroma; retrolental fibroplasia; haemangioma;
 KW Kaposi's sarcoma; neovascularisation; tumour growth.

XX Mammalia.
 OS WO200070665-A2.
 PN XX
 XX 23-NOV-2000.
 PD XX
 XX 17-MAY-2000; 2000WO-IB000763.
 PF XX
 XX 17-MAY-1999; 99US-0134406P.
 PR XX
 XX (CONJ-) CONJUCHEM INC.
 PA XX
 XX Bridon DP, Rasamoeliso M, Thibaudeau K, Huang X, Beliveau R;
 PI WPI; 2001-090970/10.
 DR XX
 XX New modified anti-angiogenic kringle 5 peptides capable of forming
 PT conjugates with blood proteins, useful for treating angiogenesis,
 PT inappropriate invasion of vessels or cancers in humans or mammals.
 XX XX
 XX Disclosure; Page 74-77; 82pp; English.
 PS XX
 XX The present invention describes a modified anti-angiogenic peptide (I)
 CC comprising a reactive group that reacts with amino groups, hydroxyl
 CC groups or thiol groups on blood components to form stable covalent bonds.
 CC The reactive group is selected from succinimidyl or maleimido groups. (I)
 CC can have anti-inflammatory, vasotrophic, cytostatic, antirheumatic,
 CC antipsoriatic, antidiabetic, antiarteriosclerotic and osteopathic
 CC activities, and is an angiogenesis inhibitor. (II) are useful for treating
 CC angiogenesis in a human, where the derivative is reacted with blood
 CC proteins. (I) are also useful for manufacturing a medicament extending
 CC the in vivo half-life of a kringle 5 peptide in a patient to provide an
 CC anti-angiogenic effect. In particular, a modified kringle 5 peptide can
 CC be used for treating inflammatory disorders (e.g. immune and non-immune
 CC inflammation, chronic articular rheumatism or psoriasis), disorders
 CC associated with inappropriate or inopportune invasion of vessels (e.g.
 CC diabetic retinopathy, neovascular glaucoma, restenosis, capillary
 CC proliferation in atherosclerotic plaques or osteoporosis), or cancer
 CC associated disorders (e.g. solid tumours, solid tumour metastases,
 CC angiofibromas, retrolental fibroplasia, haemangiomas, Kaposi's sarcoma
 CC or other cancers requiring neovascularisation to support tumour growth).
 CC The peptides are useful for treating these diseases in mammalian or human
 CC patients. AAB36562 represents a mammalian kringle 5 protein, and AAB36563
 CC to AAB36577 represent specifically claimed kringle 5 peptides from the
 CC present invention
 XX XX
 XX Sequence 790 AA;
 SQ
 Query Match 100.0%; Score 1113; DB 4; Length 790;
 Best Local Similarity 100.0%; Pred. No. 5.6e-107;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MHFCGGTLLSPFWLTAACHLEKSPRESSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60
 Db 584 MHFCGGTLLSPFWLTAACHLEKSPRESSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 643
 Qy 61 KDIALKLLSPAVITDKVIPACLPSPNVVADRTCEFTGTGQTFGAGLLKEAQLPV 120
 Db 644 KDIALKLLSPAVITDKVIPACLPSPNVVADRTCEFTGTGQTFGAGLLKEAQLPV 703
 Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
 Db 704 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 763
 Qy 181 CARPNKPGVYVRSRFTWIEGVNRN 207
 Db 764 CARPNKPGVYVRSRFTWIEGVNRN 790
 RESULT 9
 AAY50867
 ID AAY50867 standard; protein; 791 AA.

```
XX AC AAY50867;
XX DE 24-FEB-2000 (first entry)
XX DE Human plasminogen protein fragment.
XX KW Plasminogen; human; thrombolytic agent; streptokinase; antigenic;
XX KW blood clot; heart attack; treatment.
XX OS Homo sapiens.
XX PN WO9957251-A2.
XX PD 11-NOV-1999.
XX PF 06-MAY-1999; 99WO-US010086.
XX PR 06-MAY-1998; 98US-0084392P.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PI Zhang XC, Lin X, Tang JN;
XX DR WPI; 2000-052966/04.
XX PT New thrombolytic agents derived from modified humanized streptokinase,
XX PT useful for treating blood clot disorders.
XX PS Example 5; Page 40-43; 55pp; English.
XX CC This invention describes a novel thrombolytic agent comprising
XX CC streptokinase where at least one nonessential portion has been modified.
XX CC The invention also describes a method of forming a thrombolytic agent
XX CC comprising determining a nonessential portion of streptokinase and
XX CC modifying the nonessential portion to render the resulting protein less
XX CC antigenic. The modified streptokinase is used to treat blood clot
XX CC disorders, such as heart attacks. The modified streptokinase has less
XX CC antigenicity than streptokinase but is still able to complex plasminogen
XX CC and lead to plasminogen activation. Modified streptokinase with the
XX CC nonessential portions removed or truncated simplify the molecule. Such
XX CC smaller proteins are cheaper and easier to produce. This sequence
XX CC represents a fragment of the human plasminogen protein which is used in
XX CC the description of the method of the invention
XX SQ Sequence 791 AA;
XX CC Query Match 100.0%; Score 1113; DB 3; Length 791;
XX CC Best Local Similarity 100.0%; Pred. No. 5.6e-107;
XX CC Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTILSPFWLTAACHLEKSPRESSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTILSPFWLTAACHLEKSPRESSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 644
Qy 61 KDIALKLSSPAVITDKVIPACLPSPNVVADRTCEITGWTGQTFGAGLLKEAQLPV 120
Db 645 KDIALKLSSPAVITDKVIPACLPSPNVVADRTCEITGWTGQTFGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSNGLG 180
Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSNGLG 764
Qy 181 CARNKPGVYVRVSRFTWIEGVNRN 207
Db 765 CARNKPGVYVRVSRFTWIEGVNRN 791
RESULT 10
AAB01887
ID AAB01887 standard; protein; 791 AA.
XX AC AAB01887;
```


Db 645 KDIALKLLSPAVITDKVIPACLPSPNVVADRTECFITGWTGCTGTFGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSQDGGPLVCFEKDKYIILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSQDGGPLVCFEKDKYIILQGVTSWGLG 764
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRNN 207
Db 765 CARPNKPGVYVRVSRFVTWIEGVNRNN 791

RESULT 11
AAY99589
ID AAY99589 standard; protein; 791 AA.
XX
AC AAY99589;
DT 13-SEP-2000 (first entry)
XX
DE Human plasminogen.
KW Human; serine protease zymogen; cardiant; thrombolytic;
KW plasminogen activator; heart attack; stroke; blood clotting disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 542..791
FT /note= "specifically claimed in claims 7 and 8 of the
FT specification, preferably amino acids 562..791 (claim 7)"
FT Misc-difference 644..645
FT /note= "specifically claimed in claim 6 of the
FT specification"
XX
PN WO200032759-A1.
XX
PD 08-JUN-2000.
XX
XX 06-MAY-1999; 99WO-US009991.
XX
PR 02-DEC-1998; 98US-0110588P.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Lin X, Zhang XC, Tang JUN;
XX
XX WPI; 2000-422975/36.
DR
XX Polypeptide with plasminogen activator activity useful as thrombolytic
PT agent for treating blood clot disorders e.g. heart attack, comprises 10
PT amino acid peptide fragment for recognition or activation of plasminogen.
XX
PS Claim 6,7,8; Page 20-23; 41pp; English.
XX
CC The present sequence is plasminogen, the principal serine protease
CC zymogen in the extracellular fluids of vertebrates. Its active form,
CC plasmin, is implicated in pericellular proteolysis associated with a wide
CC range of physiological and pathological processes. Plasminogen expression
CC is regulated by plasminogen activators which hydrolyse a peptide bond in
CC plasminogen to convert it to plasmin or form tight binding complexes with
CC plasminogen to spontaneously convert it to plasmin. Review of sequence
CC homologues of several plasminogen activators and chymotrypsin has
CC identified a six amino acid peptide involved in plasminogen activation.
CC This peptide is particularly useful when inserted between amino acid
CC residues 644 and 645 of full length human plasminogen. Novel plasminogen
CC activators have been made based upon the plasminogen
CC activation/recognition site of plasminogen binding proteins. The
CC polypeptides are useful in preparing thrombolytic agents for treating
CC blood clotting disorders such as heart attack

Query Match 100.0%; Score 1113; DB 3; Length 791;

Best Local Similarity 100.0%; Pred. No. 5.6e-107;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLISPEWVLTAAHCLKSPSPSSYKVIILGAHQEVNLEPHVQIEVSRLFLEPTR 60
Db 585 MHFCGGTLISPEWVLTAAHCLKSPSPSSYKVIILGAHQEVNLEPHVQIEVSRLFLEPTR 644
Qy 61 KDIALKLLSPAVITDKVIPACLPSPNVVADRTECFITGWTGCTGTFGAGLLKEAQLPV 120
Db 645 KDIALKLLSPAVITDKVIPACLPSPNVVADRTECFITGWTGCTGTFGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSQDGGPLVCFEKDKYIILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSQDGGPLVCFEKDKYIILQGVTSWGLG 764
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRNN 207
Db 765 CARPNKPGVYVRVSRFVTWIEGVNRNN 791

RESULT 12
AAG67223
ID AAG67223 standard; protein; 791 AA.
XX
AC AAG67223;
DT 13-NOV-2001 (first entry)
XX
DE Amino acid sequence of human plasminogen.
XX
KW Angiotatin; plasminogen; sulfhydryl donor; angiogenesis; tumour;
KW angiogenic disease; neoplastic disease; connective tissue disorder;
KW rheumatoid arthritis; atherosclerosis; ocular angiogenic disease;
KW diabetic retinopathy; corneal graft rejection; cardiovascular disease;
KW cerebral vascular disease; diabetes; immune disorder;
KW chronic inflammation; autoimmunity.
XX
OS Homo sapiens.
XX
XX WO200158921-A2.
XX
PD 16-AUG-2001.
XX
XX 08-FEB-2001; 2001WO-US004021.
XX
XX 08-FEB-2000; 2000US-00500397.
XX
XX (NOUN) UNIV NORTHWESTERN.
XX
XX Soff G, Gately ST, Twardowski P;
XX
XX WPI; 2001-550019/61.
XX
PT Producing angiotatin for treating angiogenic diseases involves
PT contacting plasminogen with plasminogen activator and sulfhydryl donor
PT simultaneously, or producing plasmin which is contacted with sulfhydryl
PT donor.
XX
PS Disclosure; Page 77-80; 101pp; English.
XX
CC The specification describes a method for generating angiotatin in vitro.
CC The method comprises contacting plasminogen with a sulfhydryl donor, or
CC culturing cells capable of producing plasminogen activator in conditioned
CC culture medium (CCM) and contacting the CCM with plasminogen. Angiotatin
CC produced by method of the invention is useful for treating animals with
CC angiogenesis diseases. It is useful for treating an angiogenic disease
CC such as neoplastic diseases (e.g. tumours and tumour metastasis), benign
CC tumours (e.g. hemangiomas, acoustic neuromas, etc), connective tissue
CC disorders (e.g. rheumatoid arthritis and atherosclerosis), ocular
CC angiogenic diseases (e.g. diabetic retinopathy, corneal graft rejection,
CC etc), cardiovascular diseases, cerebral vascular diseases, diabetes-
CC associated diseases and immune disorders (e.g. chronic inflammation and
CC autoimmunity). The present sequence represents a plasminogen

```

XX SQ Sequence 791 AA;
Query Match 100.0%; Score 1113; DB 4; Length 791;
Best Local Similarity 100.0%; Pred. No. 5.6e-107;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPWVLTAAHCKLEKSPRPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLLSPWVLTAAHCKLEKSPRPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 644

Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 120
Db 645 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 704

Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 764

Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRNN 207
Db 765 CARPNKPGVYVRVSRFVTWIEGVNRNN 791

RESULT 13
AAE23660
ID AAE23660 standard; protein; 791 AA.
XX AC AAE23660;
XX DT 10-SEP-2002 (first entry)
XX DE Human plasminogen protein.
XX KW Kringle peptide; angiogenesis-associated disease; psoriasis; cancer;
XX KW arthritis; macular degeneration; endothelial cell proliferation;
XX KW diabetic retinopathy; cytostatic; plasminogen; human.
XX OS Homo sapiens.
XX PN WO200226782-A2.
XX PD 04-APR-2002.
XX PF 27-SEP-2001; 2001WO-US042423.
XX PR 29-SEP-2000; 2000US-00675226.
XX PR 31-AUG-2001; 2001US-00942704.
XX PA (ABBO ) ABBOTT LAB.
XX PI Henkin J, Davidson DJ;
XX DR WPI; 2002-454459/48.
XX PT Using kringle peptides conjugated to functionalized polymers, e.g.
XX PT ethoxypolyethylene, polyethylene glycol or methoxypolyethylene glycol, to
XX PT treat angiogenic disorders e.g. cancer, macular degeneration and
XX PT arthritis.
XX PS Example 1; Page 31-33; 34pp; English.
XX CC The present invention relates to conjugated kringle peptide fragments
XX CC consisting of a functionalised kringle peptide fragment chemically
XX CC coupled to a functionalised polymer. The conjugated kringle peptides may
XX CC be administered to a patient for treating angiogenesis-associated
XX CC diseases such as psoriasis, cancer, arthritis, macular degeneration,
XX CC diabetic retinopathy and for inhibiting endothelial cell proliferation.
XX CC The present sequence is human plasminogen protein. This sequence is used
XX CC in the exemplification of the invention
XX SQ Sequence 791 AA;
Query Match 100.0%; Score 1113; DB 5; Length 791;
Best Local Similarity 100.0%; Pred. No. 5.6e-107;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPWVLTAAHCKLEKSPRPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLLSPWVLTAAHCKLEKSPRPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 644

Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 120
Db 645 KDIALKLLSPAVITDKVIPACLPSPNYVADRTECFITGGETQGTGAGLLKEAQLPV 704

Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 764

Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRNN 207
Db 765 CARPNKPGVYVRVSRFVTWIEGVNRNN 791

RESULT 14
ABG76087
ID ABG76087 standard; protein; 791 AA.
XX AC ABG76087;
XX DT 13-MAY-2003 (first entry)
XX DE Human plasminogen protein fragment kringle 5.
XX KW Human; angiogenesis inhibitor; plasminogen; kringle 5; angiogenesis;
XX KW primary solid tumour; metastatic solid tumour; carcinoma; sarcoma;
XX KW lymphoma; autoimmune disease prophylaxis; rheumatoid arthritis; cancer;
XX KW immune arthritis; ocular disease; diabetic retinopathy; psoriasis;
XX KW macular degeneration; abnormal eye neovascularisation; skin disease;
XX KW blood vessel disease; haemangioma; Osler's disease; angiodioma;
XX KW capillary proliferation within atherosclerotic plaque; wound granulation;
XX KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
XX KW haemophilic joint; endothelial cell stimulation; Crohn's disease;
XX KW atherosclerosis; cat scratch disease; ulcer; placenta establishment;
XX KW ovulation inhibition; endothelial cell proliferation; birth control.
XX OS Homo sapiens.
XX PN US2002159992-A1.
XX PD 31-OCT-2002.
XX PF 28-SEP-2001; 2001US-00967386.
XX PR 29-SEP-2000; 2000US-0236550P.
XX PA (HENK/) HENKIN J.
XX PA (DAVI/) DAVIDSON D J.
XX PI Henkin J, Davidson DJ;
XX XX WPI; 2003-298673/29.
XX PT Conjugated kringle peptide fragment of plasminogen, useful for treating
XX PT angiogenic disease, e.g. cancer, comprises functionalized kringle peptide
XX PT fragment chemically coupled to functionalized polymer.
XX PS Example 1; Page 12-14; 17pp; English.
XX CC The invention relates to a conjugated kringle peptide fragment which
XX CC comprises a functionalised kringle peptide fragment chemically coupled to
XX CC a functionalised polymer. The fragment is used for treating a patient in
XX CC need of anti-angiogenic therapy, comprising administering a conjugated
XX CC kringle peptide, preferably kringles 4-5 of plasminogen. The disease can
XX CC be primary and metastatic solid tumours; carcinomas; sarcomas; lymphomas;
XX CC prophylaxis of autoimmune disease e.g. rheumatoid arthritis and immune

```

CC arthritis; ocular disease e.g. diabetic retinopathy and macular
 CC degeneration; abnormal neovascularisation conditions of the eye; skin
 CC diseases e.g. psoriasis; blood vessel disease e.g. haemangiomas and
 CC capillary proliferation within atherosclerotic plaques; Osler's disease;
 CC myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 CC haemophilic joints; angiofibroma; wound granulation; excessive or
 CC abnormal stimulation of endothelial cells e.g. Crohn's disease;
 CC atherosclerosis; cat scratch disease; ulcers. The fragment can also be
 CC used as a birth control agent which inhibits ovulation and establishment
 CC of the placenta. The fragment is also used for inhibiting endothelial
 CC cell proliferation in an individual and in vitro. The invention has an
 CC improved pharmacokinetic activity and is easily and cost-effectively
 CC made. The present sequence represents the amino acid sequence of the
 CC human plasminogen protein fragment kringle 5
 XX
 SQ Sequence 791 AA;

Query Match 100.0%; Score 1113; DB 6; Length 791;
 Best Local Similarity 100.0%; Pred. No. 5.6e-107; Gaps 0;
 Matches 207; Conservative 0; Mismatches 0; Indels 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60
 Db 585 MHFCGGTLLSPWVLTAAHCLKSPSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 644
 Qy 61 KDIALKLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGQTFGAGLLKEAQLPV 120
 Db 645 KDIALKLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGQTFGAGLLKEAQLPV 704
 Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 180
 Db 705 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 764
 Qy 181 CARPNKPGVYVRVRSRFTWIEGVNRN 207
 Db 765 CARPNKPGVYVRVRSRFTWIEGVNRN 791

RESULT 15

ADA08448
 ID ADA08448 standard; protein; 791 AA.

XX ADA08448;

XX 06-NOV-2003 (first entry)

XX Mammalian plasminogen protein.

XX Anti-angiogenesis plasmin fragment; plasminogen; plasminogen activator;
 KW plasmin reductase; reduced plasmin protein; anti-angiogenesis activator;
 KW urokinase-type plasminogen activator; A61; annexin II heterotetramer;
 KW annexin II p36 subunit; annexin II p11 subunit; thioredoxin;
 KW protein disulphide isomerase; modulating angiogenesis; mammalian.

XX Mammalia.

XX US2003083234-A1.

XX 01-MAY-2003.

XX 26-NOV-2002; 2002US-00304287.

XX 28-NOV-2001; 2001US-0333866P.

XX (WAIS//) WAISMAN D.

PA (KWON//) KWON M.

XX Waisman D, Kwon M;

XX WPI; 2003-596985/56.

XX Producing an anti-angiogenesis plasmin fragment, useful for modulating,
 PT e.g. promoting or inhibiting angiogenesis, comprises contacting a

PT plasminogen polypeptide with a plasminogen activator and a plasmin
 PT reductase.
 XX
 PS Example 1; Page 8-10; 29pp; English.
 XX
 CC The present invention relates to a method for producing an anti-
 CC angiogenesis plasmin fragment. The method comprises contacting a
 CC plasminogen polypeptide with a plasminogen activator and a plasmin
 CC reductase, where a reduced plasmin protein is produced and the anti-
 CC angiogenesis plasmin fragment having anti-angiogenesis activity, is
 CC released from the reduced plasmin protein. The plasminogen activator is
 CC preferably a urokinase-type plasminogen activator. The angiogenesis
 CC plasmin fragment is A61. The plasmin reductase is selected from annexin
 CC II heterotetramer, annexin II p36 subunit, p11, thioredoxin, and protein
 CC disulphide isomerase. The annexin II heterotetramer is associated with a
 CC cell membrane. The method of the invention is useful for modulating (e.g.
 CC promoting or inhibiting) angiogenesis. The present sequence represents
 CC mammalian plasminogen.
 XX
 SQ Sequence 791 AA;

Query Match 100.0%; Score 1113; DB 6; Length 791;
 Best Local Similarity 100.0%; Pred. No. 5.6e-107;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60
 Db 585 MHFCGGTLLSPWVLTAAHCLKSPSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 644
 Qy 61 KDIALKLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGQTFGAGLLKEAQLPV 120
 Db 645 KDIALKLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGQTFGAGLLKEAQLPV 704
 Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 180
 Db 705 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 764
 Qy 181 CARPNKPGVYVRVRSRFTWIEGVNRN 207
 Db 765 CARPNKPGVYVRVRSRFTWIEGVNRN 791

Search completed: September 10, 2004, 14:52:22
 Job time : 130 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 14:50:18 ; Search time 31 Seconds
(without alignments)
344.728 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113
Sequence: 1 MHFCGFTLSPFWLTAHC.....GVYRVSRFVTWIEGVNRNN 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1113	100.0	230	1	US-08-456-840-47
2	1113	100.0	230	1	US-08-266-407A-47
3	1113	100.0	230	2	US-08-892-544-47
4	1113	100.0	230	2	US-08-766-982-12
5	1113	100.0	230	3	US-08-944-483-53
6	1113	100.0	230	3	US-08-296-219-12
7	1113	100.0	546	6	5200340-6
8	1113	100.0	790	1	US-08-469-486-54
9	1113	100.0	790	2	US-08-469-658-54
10	1113	100.0	791	2	US-08-131-995-1
11	1113	100.0	791	2	US-08-832-087B-1
12	1113	100.0	791	3	US-08-132-154-1
13	1113	100.0	791	4	US-08-991-761A-6
14	1113	100.0	810	1	US-07-854-603-2
15	1113	100.0	810	1	US-08-147-000B-29
16	1113	100.0	810	3	US-09-086-514-1
17	1113	100.0	810	4	US-09-192-012-5
18	1113	100.0	810	4	US-09-403-736-1
19	1113	100.0	810	6	5200340-8
20	1108	99.6	791	1	US-08-643-219-1
21	1108	99.6	791	3	US-08-851-350-1
22	1106	99.4	230	1	US-08-379-621-2
23	1106	99.4	230	1	US-08-147-000B-2
24	1106	99.4	230	2	US-08-889-078-2
25	1105	99.3	814	2	US-08-750-711-1
26	1047	94.1	810	4	US-08-991-761A-11
27	1018	91.5	333	4	US-08-991-761A-8

28	946	85.0	790	4	US-08-991-761A-13	Sequence 13, Appl
29	933.5	83.9	221	3	US-08-944-483-54	Sequence 54, Appl
30	933.5	83.9	222	1	US-08-456-840-46	Sequence 46, Appl
31	933.5	83.9	222	1	US-08-266-407A-46	Sequence 46, Appl
32	933.5	83.9	222	2	US-08-892-544-46	Sequence 46, Appl
33	925	83.1	812	1	US-08-248-629A-1	Sequence 1, Appl
34	925	83.1	812	1	US-08-451-932-1	Sequence 1, Appl
35	925	83.1	812	1	US-08-452-260-1	Sequence 1, Appl
36	925	83.1	812	1	US-08-326-785-1	Sequence 1, Appl
37	925	83.1	812	2	US-08-612-788-1	Sequence 1, Appl
38	925	83.1	812	2	US-08-605-598B-1	Sequence 1, Appl
39	925	83.1	812	2	US-08-429-743-1	Sequence 1, Appl
40	925	83.1	812	2	US-08-866-735-1	Sequence 1, Appl
41	925	83.1	812	3	US-09-066-028-1	Sequence 1, Appl
42	925	83.1	812	4	US-09-192-012-3	Sequence 3, Appl
43	925	83.1	812	4	US-09-335-325-1	Sequence 1, Appl
44	925	83.1	812	4	US-08-991-761A-12	Sequence 12, Appl
45	925	83.1	812	5	PCT-US95-05107-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-456-840-47
; Sequence 47, Application US/08456840
; Patent No. 5597908
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5597908el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/456,840
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gortmley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-456-840-47

Query Match 100.0%; Score 1113; DB 1; Length 230;

	Best Local Similarity	100.0%;	Pred. No. 1.3e-115;	Mismatches	0;	Indels	0;	Gaps	0;
	Matches	207;	Conservative	0;					
Qy	1	MHFCGGTLISPEWVLTAAHCLKESPSSYKVIILGAHQEVNLEPHVQIEVSRILFLEPTR	60						
Db	24	MHFCGGTLISPEWVLTAAHCLKESPSSYKVIILGAHQEVNLEPHVQIEVSRILFLEPTR	83						
Qy	61	KDIALLLKLSPPAVITDKVIPACLLSPNNVADRTCEFTTGWGTQGTGFGAGLLKEAQLPV	120						
Db	84	KDIALLLKLSPPAVITDKVIPACLLSPNNVADRTCEFTTGWGTQGTGFGAGLLKEAQLPV	143						
Qy	121	IENKVCNRYEFLNGRVOSTELCAGHLAGTSDSCQDGGPLVCFEKKYIILQGVTSWGLG	180						
Db	144	IENKVCNRYEFLNGRVOSTELCAGHLAGTSDSCQDGGPLVCFEKKYIILQGVTSWGLG	203						
Qy	181	CARPKNKPGVYVRSRFRVTWIEGVNRNN	207						
Db	204	CARPKNKPGVYVRSRFRVTWIEGVNRNN	230						

```

RESULT 2
US 08-266-407A-47
: Sequence 47 Application US/08266407A
: Patent No. 5786156
: GENERAL INFORMATION:
: APPLICANT: Taddei-Peters, W. C.
: APPLICANT: Butler, Sandra M.
: TITLE OF INVENTION: Immunoreactive peptides of Apo(a)

```

```

Query Match      100.0%; Score 1113; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFGGTLISPEWLTAAHCKLEKSPRSSYKVTIGAHQEVNLEPHVOEIVSRLEFLEPTR 60

```

24	Db	MHFCGGTLLISPEWVLTAAHCKLEKSPRESSYKVIILGAHQEVNLEPHVQIEIVSRFLFEPTR	83
61	Qy	KDIALLKLSPPAVITDKVIPACLPSPNPNVADRTECFITGGETGTGAGLLKEAQLPV	120
84	Db	KDIALLKLSPPAVITDKVIPACLPSPNPNVADRTECFITGGETGTGAGLLKEAQLPV	143
121	Qy	IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG	180
144	Db	IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG	203
181	Qy	CARPKNKPGVTVRSRFRFTWIEGVNRNN	207
204	Db	CARPKNKPGVTVRSRFRFTWIEGVNRNN	230

```

RESULT 3
US-08-892-544-47
; Sequence 47, Application US/08892544
; Patent No. 5874544
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoreactive Peptides of A
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5874544el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892,544
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: C-terminal
; ORGANISM: Homo sapiens
; US-08-892-544-47

```

```
Query Match      100.0%; Score 1113; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy		1 MHFCGTLISPEWLTAACHLEKSPRPSYKVILGAHQEVNLEPHVQEIVSRFLFETR	60
Db		24 MHFCGTLISPEWLTAACHLEKSPRPSYKVILGAHQEVNLEPHVQEIVSRFLFETR	83
Qy		61 KOIALALKLSSPAVIDTKVIPACLPSNYVDARTCFITGWGETQGTFGALLKEAQLPV	12

Db 84 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGMGETQGTFGAGLLKEAQLPV 143
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 144 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 203
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 204 CARPNKPGVYVRVSRFVTWIEGVNRN 230

RESULT 4

US-08-766-982-12
; Sequence 12, Application US/08766982
; Patent No. 5948892
; GENERAL INFORMATION:
; APPLICANT: Wahl, Robert C.
; TITLE OF INVENTION: Analogs of Macrophage Stimulating
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Inc.
; STREET: 1840 Behavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,982
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-441
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-766-982-12

Query Match 100.0%; Score 1113; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRSPSSYKVIILGAHQEVNLEPHVQIEVSRLFLPEPTR 60
Db 24 MHFCGGTLISPEWVLTAAHCLKSPRSPSSYKVIILGAHQEVNLEPHVQIEVSRLFLPEPTR 83
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGMGETQGTFGAGLLKEAQLPV 120
Db 84 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGMGETQGTFGAGLLKEAQLPV 143
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 144 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 203
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 204 CARPNKPGVYVRVSRFVTWIEGVNRN 230

RESULT 5

US-08-944-483-53
; Sequence 53, Application US/08944483

; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183-US-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
; US-08-944-483-53

Query Match 100.0%; Score 1113; DB 3; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRSPSSYKVIILGAHQEVNLEPHVQIEVSRLFLPEPTR 60
Db 24 MHFCGGTLISPEWVLTAAHCLKSPRSPSSYKVIILGAHQEVNLEPHVQIEVSRLFLPEPTR 83
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGMGETQGTFGAGLLKEAQLPV 120
Db 84 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGMGETQGTFGAGLLKEAQLPV 143
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 144 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 203
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 204 CARPNKPGVYVRVSRFVTWIEGVNRN 230

RESULT 6
US-09-296-219-12

```
; Sequence 12, Application US/09296219
; Patent No. 6248560
; GENERAL INFORMATION:
; APPLICANT: Wahl, Robert C.
; TITLE OF INVENTION: Analogs of Macrophage Stimulating
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavenland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/296,219
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-441
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-296-219-12

Query Match 100.0%; Score 1113; DB 3; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 24 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 83
Qy 61 KDIALKLLSSPAVITDKVIPACLPSPNYVYVADRTECFITGWGETQGTGAGLLKEAQLPV 120
Db 84 KDIALKLLSSPAVITDKVIPACLPSPNYVYVADRTECFITGWGETQGTGAGLLKEAQLPV 143
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDYILQGVTSWGLG 180
Db 144 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDYILQGVTSWGLG 203
Qy 181 CARPNKPGVYVRSRFTWIEGVNRNN 207
Db 204 CARPNKPGVYVRSRFTWIEGVNRNN 230

RESULT 7
5200340-6
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO: 6:
; LENGTH: 546
; 5200340-6

Query Match 100.0%; Score 1113; DB 6; Length 546;
Best Local Similarity 100.0%; Pred. No. 4.2e-115;

Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 340 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 399
Qy 61 KDIALKLLSSPAVITDKVIPACLPSPNYVYVADRTECFITGWGETQGTGAGLLKEAQLPV 120
Db 400 KDIALKLLSSPAVITDKVIPACLPSPNYVYVADRTECFITGWGETQGTGAGLLKEAQLPV 459
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDYILQGVTSWGLG 180
Db 460 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDYILQGVTSWGLG 519
Qy 181 CARPNKPGVYVRSRFTWIEGVNRNN 207
Db 520 CARPNKPGVYVRSRFTWIEGVNRNN 546

RESULT 8
US-08-469-486-54
; Sequence 54, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-486-54

Query Match 100.0%; Score 1113; DB 1; Length 790;
Best Local Similarity 100.0%; Pred. No. 7e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
```


Db 584 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEIVSRFLFLEPTR 643
Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKEAQLPV 120
Db 644 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKEAQLPV 703
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCGDSGGPLVCFEKKYILOGVTSWGLG 180
Db 704 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCGDSGGPLVCFEKKYILOGVTSWGLG 763
Qy 181 CARENKPGVYVVRVSRFVTWIEGVNRNN 207
Db 764 CARENKPGVYVVRVSRFVTWIEGVNRNN 790

RESULT 9

US-08-469-658-54
; Sequence 54, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Th egeresen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF
; PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-658-54

Query Match 100.0%; Score 1113; DB 2; Length 790;
Best Local Similarity 100.0%; Pred. No. 7e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEIVSRFLFLEPTR 60
Db 584 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEIVSRFLFLEPTR 643
Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKEAQLPV 120

Db 644 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKEAQLPV 703
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCGDSGGPLVCFEKKYILOGVTSWGLG 180
Db 704 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCGDSGGPLVCFEKKYILOGVTSWGLG 763
Qy 181 CARENKPGVYVVRVSRFVTWIEGVNRNN 207
Db 764 CARENKPGVYVVRVSRFVTWIEGVNRNN 790

RESULT 10

US-09-131-995-1
; Sequence 1, Application US/09131995
; Patent No. 5972896
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; AND METHODS FOR INHIBITING ANGIOGENESIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/131,995
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,087
; FILING DATE: 03-APR-1997
; APPLICATION NUMBER: 08/643,219
; FILING DATE: 06-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 5940. US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 791 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: protein
US-09-131-995-1

Query Match 100.0%; Score 1113; DB 2; Length 791;
Best Local Similarity 100.0%; Pred. No. 7e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEIVSRFLFLEPTR 60
Db 585 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEIVSRFLFLEPTR 644
Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKEAQLPV 120
Db 645 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEFTGWGETQGTGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCGDSGGPLVCFEKKYILOGVTSWGLG 180

Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDSGGPLVCPFKDKYILQGVTSWGLG 764
Qy 181 CARENKPGVYVRSRFFVTWIEGVNRN 207
Db 765 CARENKPGVYVRSRFFVTWIEGVNRN 791

RESULT 11

US-08-832-087B-1
; Sequence 1, Application US/08832087B
; Patent No. 5981484
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,087B
; FILING DATE: 03-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/643,219
; FILING DATE: 06-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 5940.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 791 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; US-08-832-087B-1

Query Match 100.0%; Score 1113; DB 2; Length 791;
Best Local Similarity 100.0%; Pred. No. 7e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVLGAHQEVLNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVLGAHQEVLNLEPHVQIEVSRFLFLEPTR 644
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVVDRTCEFTICGGETQGTGAGLLKEAQLPV 120
Db 645 KDIALKLSPPAVITDKVIPACLPSPNYVVDRTCEFTICGGETQGTGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDSGGPLVCPFKDKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDSGGPLVCPFKDKYILQGVTSWGLG 764

Qy 181 CARENKPGVYVRSRFFVTWIEGVNRN 207
Db 765 CARENKPGVYVRSRFFVTWIEGVNRN 791

RESULT 12

US-09-132-154-1
; Sequence 1, Application US/09132154
; Patent No. 6251867
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/132,154
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,087
; FILING DATE: 03-APR-1997
; APPLICATION NUMBER: 08/643,219
; FILING DATE: 06-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 5940.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 791 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; US-09-132-154-1

Query Match 100.0%; Score 1113; DB 3; Length 791;
Best Local Similarity 100.0%; Pred. No. 7e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVLGAHQEVLNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVLGAHQEVLNLEPHVQIEVSRFLFLEPTR 644
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVVDRTCEFTICGGETQGTGAGLLKEAQLPV 120
Db 645 KDIALKLSPPAVITDKVIPACLPSPNYVVDRTCEFTICGGETQGTGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDSGGPLVCPFKDKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDSGGPLVCPFKDKYILQGVTSWGLG 764
Qy 181 CARENKPGVYVRSRFFVTWIEGVNRN 207
Db 765 CARENKPGVYVRSRFFVTWIEGVNRN 791

RESULT 13

US-08-991-761A-6
; Sequence 6, Application US/08991761A
; Patent No. 6576609
; GENERAL INFORMATION:

APPLICANT: Soff, Gerald
APPLICANT: Gately, Stephen
APPLICANT: Twardowski, Przemyslaw
TITLE OF INVENTION: "Methods and Compositions for Generating
TITLE OF INVENTION: Angiostatin"
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: CO
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991.761A
FILING DATE:
CLASSIFICATION: 1642
ATTORNEY/AGENT INFORMATION:
NAME: Crook, Wainell M.
REGISTRATION NUMBER: 31,071
REFERENCE/DOCKET NUMBER: 3501-16-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 791 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-761A-6
Query Match 100.0%; Score 1113; DB 4; Length 791;
Best Local Similarity 100.0%; Pred. No. 7e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 644
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTCEFTTGWGETQGTGTFGAGLLKEAQLPV 120
Db 645 KDIALKLSPPAVITDKVIPACLPSPNYVADRTCEFTTGWGETQGTGTFGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEPLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYIILQGVTSWGLG 180
Db 705 IENKVCNRYEPLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYIILQGVTSWGLG 764
Qy 181 CARPKPGVYVRVSRFVTWIEGVNRN 207
Db 765 CARPKPGVYVRVSRFVTWIEGVNRN 791
RESULT 14
US-07-854-603-2
Sequence 2, Application US/07854603
Patent No. 5637492
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Edwards, Richard M
APPLICANT: Forman, Joan M
TITLE OF INVENTION: Activatable fibrinolytic and
TITLE OF INVENTION: anti-thrombotic proteins
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000

CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854.603
FILING DATE: 19901207
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,338
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 810 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-854-603-2
Query Match 100.0%; Score 1113; DB 1; Length 810;
Best Local Similarity 100.0%; Pred. No. 7.2e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 604 MHFCGGTLISPEWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 663
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTCEFTTGWGETQGTGTFGAGLLKEAQLPV 120
Db 664 KDIALKLSPPAVITDKVIPACLPSPNYVADRTCEFTTGWGETQGTGTFGAGLLKEAQLPV 723
Qy 121 IENKVCNRYEPLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYIILQGVTSWGLG 180
Db 724 IENKVCNRYEPLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKDKYIILQGVTSWGLG 783
Qy 181 CARPKPGVYVRVSRFVTWIEGVNRN 207
Db 784 CARPKPGVYVRVSRFVTWIEGVNRN 810
RESULT 15
US-08-147-000B-29
Sequence 29, Application US/08147000B
Patent No. 5688664
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Gilbert, Richard J
TITLE OF INVENTION: THROMBIN ACTIVATABLE PLASMINOGEN ANALOGUES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE:
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/147,000B
FILING DATE: October 29, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,603
FILING DATE: June 4, 1992
APPLICATION NUMBER: GB 92 22758.6
FILING DATE: October 29, 1992
APPLICATION NUMBER: PCT/GB90/01912
FILING DATE: December 7, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoscheit, Dale H.
REGISTRATION NUMBER: 19,090
REFERENCE/DOCKET NUMBER: 10180.60948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9200
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 810 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-147-000B-29

Query Match 100.0%; Score 1113; DB 1; Length 810;
Best Local Similarity 100.0%; Pred. No. 7.2e-115; Indels 0; Gaps 0;
Matches 207; Conservative 0; Mismatches 0;
Qy 1 MHFCGGTLISPEWVLTAAHCLKSPSPSSYKVLGAHQEVNLEPHVQEIIVSRLEPTR 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
604 MHFCGGTLISPEWVLTAAHCLKSPSPSSYKVLGAHQEVNLEPHVQEIIVSRLEPTR 663
Qy 61 KDIALKLSSPAVITDKVIPACLSPNPNVADRTECFITGNGETQGTGAGLLKEAQLPV 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
664 KDIALKLSSPAVITDKVIPACLSPNPNVADRTECFITGNGETQGTGAGLLKEAQLPV 723
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
724 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLG 783
Qy 181 CARENKPQVYVRSRFTWIEGVNRNN 207
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
784 CARENKPQVYVRSRFTWIEGVNRNN 810

Search completed: September 10, 2004, 14:56:26
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 14:55:15 ; Search time 127 Seconds
(without alignments)
522.699 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113

Sequence: 1 MHFCGGTLLSPWVLTAAHC.....GVYRVSRFTWIEGVNRNN 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES.

Result No.	Score	Query Match	Length	ID	Description
1	1113	100.0	207	10	US-09-992-600A-54
2	1113	100.0	207	10	US-09-924-340-54
3	1113	100.0	207	10	US-09-992-095B-54
4	1113	100.0	207	14	US-09-999-570-54
5	1113	100.0	207	14	US-10-000-489-54
6	1113	100.0	207	14	US-10-000-986-54
7	1113	100.0	207	14	US-10-154-678-54
8	1113	100.0	249	12	US-10-450-976-4
9	1113	100.0	348	12	US-10-450-976-6
10	1113	100.0	714	16	US-10-415-012-8
11	1113	100.0	791	9	US-09-967-386-1
12	1113	100.0	791	14	US-10-304-287-1
13	1113	100.0	791	15	US-10-360-101-257
14	1113	100.0	791	16	US-10-778-423-1
15	1113	100.0	791	16	US-10-753-646-1

16	1113	100.0	791	16	US-10-735-577-1	Sequence 1, Appli
17	1113	100.0	810	9	US-09-946-893-2	Sequence 2, Appli
18	1113	100.0	810	12	US-10-135-872B-4	Sequence 4, Appli
19	1113	100.0	810	12	US-10-450-976-2	Sequence 2, Appli
20	1113	100.0	810	14	US-10-193-656-2	Sequence 2, Appli
21	1113	100.0	810	14	US-10-237-144-1	Sequence 1, Appli
22	1113	100.0	810	16	US-10-415-012-4	Sequence 4, Appli
23	1113	100.0	810	16	US-10-741-601-409	Sequence 409, App
24	946	85.0	790	12	US-09-825-751A-70	Sequence 70, Appl
25	933.5	83.9	1169	9	US-09-870-759-126	Sequence 126, App
26	933.5	83.9	1169	10	US-09-751-708A-126	Sequence 126, App
27	925	83.1	812	9	US-09-788-142-1	Sequence 1, Appli
28	925	83.1	812	9	US-09-761-120-1	Sequence 1, Appli
29	925	83.1	812	9	US-09-873-676-81	Sequence 81, Appli
30	925	83.1	812	9	US-09-335-325-1	Sequence 1, Appli
31	925	83.1	812	12	US-10-127-066-1	Sequence 1, Appli
32	925	83.1	812	14	US-10-131-241-1	Sequence 1, Appli
33	925	83.1	812	15	US-10-402-364-1	Sequence 1, Appli
34	925	83.1	812	16	US-10-401-108-1	Sequence 1, Appli
35	878	78.9	812	12	US-09-825-751A-71	Sequence 71, Appl
36	496	44.6	229	15	US-10-051-874-101	Sequence 101, App
37	496	44.6	230	11	US-09-981-151A-87	Sequence 87, Appl
38	496	44.6	230	11	US-09-981-151A-96	Sequence 96, Appl
39	496	44.6	230	12	US-10-042-865-155	Sequence 155, App
40	496	44.6	230	12	US-10-072-012-804	Sequence 804, App
41	496	44.6	230	12	US-10-072-012-812	Sequence 812, App
42	496	44.6	230	12	US-10-037-417-135	Sequence 135, App
43	496	44.6	230	14	US-10-032-189-66	Sequence 66, Appl
44	496	44.6	230	15	US-10-074-978A-221	Sequence 221, App
45	496	44.6	230	15	US-10-074-978A-222	Sequence 222, App

ALIGNMENTS

RESULT 1

US-09-992-600A-54
; Sequence 54, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992, 600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Jpatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-600A-54

Query Match 100.0%; Score 1113; DB 10; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.1e-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLLSPWVLTAAHCLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRFLFEPTR 60
Db 1 MHFCGGTLLSPWVLTAAHCLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRFLFEPTR 60

Best Local Similarity 100.0%; Pred. No. 2.le-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Qy 61 KDIALKLLSSPAVITDKVIPACLSPSNVYVADRTECFITGWTGQTFGAGLLKEAQLPV 120
Db 61 KDIALKLLSSPAVITDKVIPACLSPSNVYVADRTECFITGWTGQTFGAGLLKEAQLPV 120
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Db 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207

RESULT 5

US-10-000-489-54
; Sequence 54, Application US/10000489
; Publication No. US2003092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-489-54

Query Match 100.0%; Score 1113; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.le-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Qy 61 KDIALKLLSSPAVITDKVIPACLSPSNVYVADRTECFITGWTGQTFGAGLLKEAQLPV 120
Db 61 KDIALKLLSSPAVITDKVIPACLSPSNVYVADRTECFITGWTGQTFGAGLLKEAQLPV 120
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Db 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207

RESULT 6

US-10-000-986-54
; Sequence 54, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US9.DIV
; CURRENT APPLICATION NUMBER: US/10/000,986
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-986-54

Query Match 100.0%; Score 1113; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.le-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 1 MHFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Qy 61 KDIALKLLSSPAVITDKVIPACLSPSNVYVADRTECFITGWTGQTFGAGLLKEAQLPV 120
Db 61 KDIALKLLSSPAVITDKVIPACLSPSNVYVADRTECFITGWTGQTFGAGLLKEAQLPV 120
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Db 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 181 CARPNKPGVYVRVSRFVTWIEGVNRN 207

RESULT 7

US-10-154-678-54
; Sequence 54, Application US/10154678
; Publication No. US20030162186A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182.US1.REG
; CURRENT APPLICATION NUMBER: US/10/154,678
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112

```
; SOFTWARE: JPatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-678-54

Query Match      100.0%; Score 1113; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.1e-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60
Db 1 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60

Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 120
Db 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 120

Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180

Qy 181 CARENKPGVYVRSRFTVWIEGVNRN 207
Db 181 CARENKPGVYVRSRFTVWIEGVNRN 207

RESULT 8
US-10-450-976-4
; Sequence 4, Application US/10450976
; Publication No. US20040071676A1
; GENERAL INFORMATION:
; APPLICANT: COLLEN, Desire Jose
; APPLICANT: NAGAI, Nubuo
; APPLICANT: LAROCHE, Yves
; TITLE OF INVENTION: A Yeast Expression Vector and a Method
; TITLE OF INVENTION: of Making a Recombinant Protein by Expression in a Yeast
; FILE REFERENCE: 50304/005001
; CURRENT APPLICATION NUMBER: US/10/450,976
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/BE01/00217
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: GB 0116702
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0116690
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0031196
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-976-4

Query Match      100.0%; Score 1113; DB 12; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.7e-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60
Db 43 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 102

Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 120
Db 103 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 162

Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 163 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 222

; SOFTWARE: JPatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-678-54

Query Match      100.0%; Score 1113; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.1e-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60
Db 1 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60

Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 120
Db 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 120

Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180

Qy 181 CARENKPGVYVRSRFTVWIEGVNRN 207
Db 181 CARENKPGVYVRSRFTVWIEGVNRN 207

RESULT 9
US-10-450-976-6
; Sequence 6, Application US/10450976
; Publication No. US20040071676A1
; GENERAL INFORMATION:
; APPLICANT: COLLEN, Desire Jose
; APPLICANT: NAGAI, Nubuo
; APPLICANT: LAROCHE, Yves
; TITLE OF INVENTION: A Yeast Expression Vector and a Method
; TITLE OF INVENTION: of Making a Recombinant Protein by Expression in a Yeast
; FILE REFERENCE: 50304/005001
; CURRENT APPLICATION NUMBER: US/10/450,976
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/BE01/00217
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: GB 0116702
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0116690
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0031196
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-976-6

Query Match      100.0%; Score 1113; DB 12; Length 348;
Best Local Similarity 100.0%; Pred. No. 4.2e-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60
Db 142 MHFCGGTLLSPDWLTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 201

Qy 61 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 120
Db 202 KDIALKLLSPAVITDKVIPACLPSPNYVADRTCEITGWTGQTFGAGLLKEAQLPV 261

Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
Db 262 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 321

Qy 181 CARENKPGVYVRSRFTVWIEGVNRN 207
Db 322 CARENKPGVYVRSRFTVWIEGVNRN 348

RESULT 10
US-10-415-012-8
; Sequence 8, Application US/10415012
; Publication No. US20040082030A1
; GENERAL INFORMATION:
; APPLICANT: Waisman, David M
; TITLE OF INVENTION: Anti-angiogenic polypeptides
; FILE REFERENCE: ME03-002
; CURRENT APPLICATION NUMBER: US/10/415,012
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US 60/253725
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; LENGTH: 714
```


; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-012-8

Query Match 100.0%; Score 1113; DB 16; Length 714;
Best Local Similarity 100.0%; Pred. No. 1.3e-112; Indels 0; Gaps 0;
Matches 207; Conservative 0; Mismatches 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 508 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 567
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 120
Db 568 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 627
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Db 628 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 687
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRNN 207
Db 688 CARPNKPGVYVRVSRFVTWIEGVNRNN 714

RESULT 11
US-09-967-386-1
; Sequence 1, Application US/09967386
; Patent No. US20020159952A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: ANTIANGIOGENIC POLYPEPTIDES AND METHODS
; FILE REFERENCE: 6738.US.02
; CURRENT APPLICATION NUMBER: US/09/967,386
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/236,550
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-967-386-1

Query Match 100.0%; Score 1113; DB 9; Length 791;
Best Local Similarity 100.0%; Pred. No. 1.3e-112; Indels 0; Gaps 0;
Matches 207; Conservative 0; Mismatches 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 644
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 120
Db 645 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 764
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRNN 207
Db 765 CARPNKPGVYVRVSRFVTWIEGVNRNN 791

RESULT 12
US-10-304-287-1
; Sequence 1, Application US/10304287
; Publication No. US20030081234A1

GENERAL INFORMATION:

; APPLICANT: Waisman, David M.
; APPLICANT: Kwon, Mi Jung
; TITLE OF INVENTION: Anti-angiogenesis methods, compositions and uses therefor
; FILE REFERENCE: ME02-001
; CURRENT APPLICATION NUMBER: US/10/304,287
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US 60/333,866
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Microsoft Word
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: mammalian
US-10-304-287-1

Query Match 100.0%; Score 1113; DB 14; Length 791;
Best Local Similarity 100.0%; Pred. No. 1.3e-112; Indels 0; Gaps 0;
Matches 207; Conservative 0; Mismatches 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 644
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 120
Db 645 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKKDKYILQGVTSWGLG 764
Qy 181 CARPNKPGVYVRVSRFVTWIEGVNRNN 207
Db 765 CARPNKPGVYVRVSRFVTWIEGVNRNN 791

RESULT 13

US-10-360-101-257
; Sequence 257, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 257
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of plasminogen
US-10-360-101-257

Query Match 100.0%; Score 1113; DB 15; Length 791;
Best Local Similarity 100.0%; Pred. No. 1.3e-112; Indels 0; Gaps 0;
Matches 207; Conservative 0; Mismatches 0;

Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 585 MHFCGGTLLSPWVLTAAHCLKSPSSYKVIILGAHQEVNLEPHVQIEVSRFLFLEPTR 644
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 120
Db 645 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWGTQGTGAGLLKEAQLPV 704

Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKYILQGVTSWGLG 764
Qy 181 CARPNKPGVYVRSRFTWIEGVNRNN 207
Db 765 CARPNKPGVYVRSRFTWIEGVNRNN 791

RESULT 14

US-10-778-423-1
; Sequence 1, Application US/10778423
; Publication No. US20040132664A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: ANTIANGIOGENIC POLYPEPTIDES AND METHODS
; FILE OF INVENTION: FOR INHIBITING ANGIOGENESIS
; FILE REFERENCE: 6738 US 02
; CURRENT APPLICATION NUMBER: US/10/778,423
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: US/09/967,386
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,550
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-778-423-1

Query Match 100.0%; Score 1113; DB 16; Length 791;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRESSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60
Db 585 MHFCGGTLISPEWVLTAAHCLKSPRESSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 644
Qy 61 KDIALKLSSPAVITDKVIPACLPSPNYVADRTECFITGWGETQGTGAGLLKEAQLPV 120
Db 645 KDIALKLSSPAVITDKVIPACLPSPNYVADRTECFITGWGETQGTGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKYILQGVTSWGLG 764
Qy 181 CARPNKPGVYVRSRFTWIEGVNRNN 207
Db 765 CARPNKPGVYVRSRFTWIEGVNRNN 791

RESULT 15

US-10-753-646-1
; Sequence 1, Application US/10753646
; Publication No. US20040138127A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,
; FILE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS FOR INHIBITING
; FILE REFERENCE: 5940.US.P3
; CURRENT APPLICATION NUMBER: US/10/753,646
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US/08/924,287A
; PRIOR FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 08/851,350
; PRIOR FILING DATE: 1997-05-05
; PRIOR APPLICATION NUMBER: US 08/832,087

; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: US 08/643,219
; PRIOR FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-753-646-1

Query Match 100.0%; Score 1113; DB 16; Length 791;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHFCGGTLISPEWVLTAAHCLKSPRESSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 60
Db 585 MHFCGGTLISPEWVLTAAHCLKSPRESSYKVLGAHQEVNLEPHVQIEVSRLFLEPTR 644
Qy 61 KDIALKLSSPAVITDKVIPACLPSPNYVADRTECFITGWGETQGTGAGLLKEAQLPV 120
Db 645 KDIALKLSSPAVITDKVIPACLPSPNYVADRTECFITGWGETQGTGAGLLKEAQLPV 704
Qy 121 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKYILQGVTSWGLG 180
Db 705 IENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGPLVCFEKKYILQGVTSWGLG 764
Qy 181 CARPNKPGVYVRSRFTWIEGVNRNN 207
Db 765 CARPNKPGVYVRSRFTWIEGVNRNN 791

Search completed: September 10, 2004, 15:06:54
Job time : 128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 14:46:43 ; Search time 120 Seconds
(without alignments)
544.269 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113
Sequence: 1 MHFGGTLISPEWVLTAAHC.....GVYVRVSRFTWIEGVNRNN 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1112	99.9	810	4	Q15146
2	1060	95.2	334	6	O46507
3	993	89.2	454	6	O46506
4	907	81.5	812	11	Q9R0M3
5	849	76.3	806	6	O18783
6	719	64.6	429	13	Q8AVB0
7	464.5	41.7	310	11	Q9QY29
8	464.5	41.7	310	11	Q91XC4
9	445.5	40.0	453	11	Q812A6
10	440	39.5	317	13	Q9DGR3
11	437	39.3	277	11	Q80WM7
12	437	39.3	284	4	Q8NF86
13	434.5	39.0	267	5	Q8BK47
14	433	38.9	327	4	Q8N171
15	432.5	38.9	581	5	Q9XZM7
16	431	38.7	328	11	Q8BJR6

ALIGNMENTS

RESULT 1

ID	Q15146	PRELIMINARY;	PRT;	810 AA.
AC	Q15146;			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Plasminogen precursor.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,			
RA	Mitchell D., Robinson J.H.;			
RT	"Expression of recombinant human plasminogen and aglycoplasminogen in			
RT	HeLa cells";			
RL	Fibrinolysis 0:0-0(1991).			
CC	-1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.			
DR	EMBL; M74220; AAA36451.1; -;			
DR	HSSP; P00747; 2PK4.			
DR	GO; GO:0005509; F:calcium ion binding; IEA.			
DR	GO; GO:0004263; F:chymotrypsin activity; IEA.			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	GO; GO:0003809; F:thrombin activity; IEA.			
DR	GO; GO:0004295; F:trypsin activity; IEA.			
DR	GO; GO:0007596; P:blood coagulation; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR009003; Cys_Ser_trypsin.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR03014; PAN.			
DR	InterPro; IPR003609; Pan app.			
DR	InterPro; IPR001254; Peptidase_S1.			
DR	InterPro; IPR001314; Peptidase_S1A.			
DR	InterPro; IPR003966; Peptidase_S1A_pr.			
DR	Pfam; PF00051; Kringle; 5.			
DR	Pfam; PF00024; PAN; 1.			

17	424.5	38.1	336	11	Q80YD8
18	424	38.1	680	5	Q868H5
19	420	37.7	261	13	Q9W7Q4
20	419.5	37.7	328	11	Q80240
21	419	37.6	273	11	Q921N4
22	417	37.5	321	4	Q96RZ8
23	417	37.5	680	5	Q868H7
24	416.5	37.4	471	11	Q8CFE0
25	415.5	37.3	558	4	Q86YM4
26	414.5	37.2	540	13	Q800Y7
27	413.5	37.2	537	4	Q9BYE1
28	413	37.1	777	11	Q8CAN9
29	412.5	37.1	581	4	Q9BYE2
30	412	37.0	271	13	Q803Z4
31	410.5	36.9	1524	13	Q91674
32	410	36.8	245	13	Q42160
33	409.5	36.8	331	11	Q8RIA6
34	409.5	36.8	331	11	Q80X17
35	409	36.7	242	13	Q7SAX0
36	409	36.7	244	13	Q42159
37	408.5	36.7	490	11	Q920K3
38	408	36.7	638	11	Q8ROP5
39	407.5	36.6	490	11	Q7TN04
40	406.5	36.5	371	11	Q8CJ16
41	406.5	36.5	445	11	Q8CU17
42	406	36.5	415	6	Q29015
43	405	36.4	311	11	Q80XZ3
44	404	36.3	249	13	Q9W7Q1
45	403	36.2	266	13	Q92077

Q80YD8 mus musculus
Q868H5 brachiolesto
Q9W7Q4 parolichthy
Q80240 rattus norv
Q921N4 mus musculus
Q96RZ8 homo sapien
Q868H7 brachiolesto
Q8CFE0 mus musculus
Q86YM4 homo sapien
Q800Y7 meleagris g
Q9BYE1 homo sapien
Q8CAN9 mus musculus
Q9BYE2 homo sapien
Q803Z4 brachydanio
Q91674 xenopus lae
Q42160 petromyzon
Q8RIA6 mus musculus
Q80X17 mus musculus
Q7SAX0 brachydanio
Q42159 petromyzon
Q920K3 rattus norv
Q8ROP5 mus musculus
Q7TN04 mus musculus
Q8CJ16 rattus norv
Q8CU17 rattus norv
Q29015 sus sp. pre
Q80XZ3 rattus norv
Q9W7Q1 parolichthy
Q92077 gadus morhu

DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR0018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR PRODOM; PD000395; Kringle; 5.
 DR SMART; SM00130; KR; 5.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; TRYD_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 5.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 810 PLASMINOGEN.
 SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;
 Query Match 99.9%; Score 1112; DB 4; Length 810;
 Best Local Similarity 99.5%; Pred. No. 3.9e-105; Mismatches 0; Indels 0; Gaps 0;
 Matches 206; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MHFCGGTLLSPWVLTAAHCLKSPRPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
 DB 604 MHFCGGTLLSPWVLTAAHCLKSPRPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 663
 QY 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWTGQTFGAGLLKEAOLPV 120
 DB 664 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWTGQTFGAGLLKEAOLPV 723
 QY 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
 DB 724 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 783
 QY 181 CARPNKPGVYVRSRVFTWIEGVNRN 207
 DB 784 CARPNKPGVYVRSRVFTWIEGVNRN 810
 RESULT 2
 ID O46507 PRELIMINARY; PRT; 334 AA.
 AC O46507;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Plasminogen (fragment).
 GN BABEPFSG.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Papio.
 OX NCBI_TaxID=9557;
 RN [1]
 RC TISSUE=Liver;
 RA Cox L.A., Jett C., Hixson J.E.;
 RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele."
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AF029692; AAB97887.1; -.
 DR HSP; P00747; SHPG.
 DR MEROPS; S01.233; -.
 DR GO; GO:000509; F:calcium ion binding; IEA.
 DR GO; GO:0004263; F:peptidase activity; IEA.
 DR GO; GO:0008233; F:chymotrypsin activity; IEA.
 DR GO; GO:0003809; F:thrombin activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0007596; F:blood coagulation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.

DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase S1.
 DR InterPro; IPR001314; Peptidase S1A.
 DR InterPro; IPR003966; Peptidase_S1A_pr.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR0018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYD_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
 FT NON_TER 1 1
 SQ SEQUENCE 334 AA; 36791 MW; C7DC06E03B965286 CRC64;
 Query Match 95.2%; Score 1060; DB 6; Length 334;
 Best Local Similarity 94.2%; Pred. No. 2.7e-100; Mismatches 6; Indels 0; Gaps 0;
 Matches 195; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MHFCGGTLLSPWVLTAAHCLKSPRPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60
 DB 128 MHFCGGTLLSPWVLTAAHCLKSPRPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 187
 QY 61 KDIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWTGQTFGAGLLKEAOLPV 120
 DB 188 ADIALKLSPPAVITDKVIPACLPSPNYVADRTECFITGWTGQTFGAGLLKEAOLPV 247
 QY 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180
 DB 248 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 307
 QY 181 CARPNKPGVYVRSRVFTWIEGVNRN 207
 DB 308 CARPNKPGVYVRSRVFTWIEGVNRN 334
 RESULT 3
 ID O46506 PRELIMINARY; PRT; 454 AA.
 AC O46506;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Apolipoprotein a (fragment).
 GN BABAPOA.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Papio.
 OX NCBI_TaxID=9557;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Cox L.A., Jett C., Hixson J.E.;
 RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele."
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; AF029691; AAB97886.1; -.
 DR HSP; P00747; 2PK4.
 DR MEROPS; S01.999; -.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.

```
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON TER 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 89.2%; Score 993; DB 6; Length 454;
Best Local Similarity 88.3%; Pred. No. 3.1e-93;
Matches 182; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRLEPTRK 61
Db 249 HFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRLEPTRK 61
Qy 62 DIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTTGMGETGTGAGLLKEAQLPVI 121
Db 309 DIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTTGMGETGTGAGLLKEAQLPVI 121
Qy 122 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCGSGPLVCFEKKYILQGVTSWGLGC 181
Db 369 ENTVCNRYEFLNGRVQSTELCAGHLAGTDSQCGSGPLVCFEKKYILQGVTSWGLGC 181
Qy 182 ARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 429 ARPNKPGVYVRVSRFVTWIEGVNRN 454

RESULT 4
Q9ROW3 Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP TISSUE=Liver;
RC TISSUE=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanatas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RL J. Biol. Chem. 266:10825-10829 (1991).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AJ242649; CAB46014.1; -.
DR HSP; P00747; 1PK.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004283; F:plasmin activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR00001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 19
FT CHAIN 20 812 PLASMINOGEN.
SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 81.5%; Score 907; DB 11; Length 812;
Best Local Similarity 83.0%; Pred. No. 4.3e-84;
Matches 171; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

Qy 2 HFCGGTLLSPWVLTAAHCLKSPRSSYKVIILGAHQEVNLEPHVQIEVSRLEPTRK 61
Db 607 HFCGGTLLSPWVLTAAHCLKSPREFYKVIILGAHEERILGSDVQGIATVKLVLEPDA 666
Qy 62 DIALLKLSPPAVITDKVIPACLPSPNVVADRTCEFTTGMGETGTGAGLLKEAQLPVI 121
Db 667 DIALLKLSPPAVITDKVIPACLPSPNVVADRTLCYTITGGETGTGAGLLKEAQLPVI 726
Qy 122 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCGSGPLVCFEKKYILQGVTSWGLGC 181
Db 727 ENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCGSGPLVCFEKKYILQGVTSWGLGC 786
Qy 182 ARPNKPGVYVRVSRFVTWIEGVNRN 207
Db 787 ARPNKPGVYVRVSRFVTWIEGVNRN 812

RESULT 5
O18783 O18783 PRELIMINARY; PRT; 806 AA.
AC O18783;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxId=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
```

```

RA Lawn R.M., Schwartz K., Patthy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL Proc. Natl. Acad. Sci. U.S.A. 94,11992-11997(1997).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AF012297; AAB65760.1; -.
DR HSSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR005003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS50070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

Query Match 76.3%; Score 849; DB 6; Length 806;
Best Local Similarity 74.3%; Pred. No. 3.8e-78;
Matches 153; Conservative 24; Mismatches 29; Indels 0; Gaps 0;

Qy 2 HFCGGTLISPEWLVTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 61
Db 601 HFCGGTLIAQWLVTAACHLEKRSQWPGAYKVLGHRVAPESYSQSIGVSRFLKGPLAA 660
Qy 62 DIALLKLSPPAVITDKVIPACLPSPNTYVADRTCEFTIGWGETQGTGAGLLKEAQLPVI 121
Db 661 DIALLKLRPAINDKVIPACLPSPQDFWVPDRTLCHVTGWGDTQGTSPRGLLKQASLPVI 720
Qy 122 ENKVCNRYEFLNGRVQSTELCAGHLAGTSDCGSDGGLVCFKDKYILQGVTSWGLGC 181
Db 721 DNRVCNREHYLNGRVKSTELCAGHLVGRGSDCGSDGGLICFEDDKYVLQGVTSWGLGC 780
Qy 182 ARPNKPGVYVRVSRFVTWIEGWENN 207
Db 781 ARPNKPGVYVRVSRFVTWIEGWENN 806

RESULT 6
Q8AVB0 PRELIMINARY; PRT; 429 AA.
AC Q8AVB0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen precursor (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_TaxID=7955;

```

```

RN SEQUENCE FROM N.A.
RP Hanumanthaiah R., Day K., Jagadeeswaran P.;
RT "Comprehensive analysis of blood coagulation pathways in teleostei:
RT Evolution of coagulation factor genes and identification of zebrafish
RT factor VIII.";
RL Blood Cells Mol. Dis. 0:0-0(2002).
DR EMBL; AF515276; AAN71006.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
FT NON TER 1
SQ SEQUENCE 429 AA; 47556 MW; 9A580A214A549C12 CRC64;

Query Match 64.8%; Score 719; DB 13; Length 429;
Best Local Similarity 61.8%; Pred. No. 3.8e-65;
Matches 128; Conservative 28; Mismatches 51; Indels 0; Gaps 0;

Qy 1 MHECGGTLISPEWLVTAACHLEKSPRSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60
Db 223 IHFCGGTLIDPQWLVTAACHLEKSDSPSAKIMLGHTERATSSKQERDVTKIKGPAG 282
Qy 61 KDIALKLSPPAVITDKVIPACLPSPNTYVADRTCEFTIGWGETQGTGAGLLKEAQLPV 120
Db 283 TDIALKLRPALINDKVPCLPEKDYIVPSNTECVVTGWTGQDTQGTGGGYLKETGFPV 342
Qy 121 ENKVCNRYEFLNGRVQSTELCAGHLAGTSDCGSDGGLVCFKDKYILQGVTSWGLG 180
Db 343 IENKVCNRPFLNGRVKDHMCAGNIEGNDSCQSDGGLVCAQNTFVLOGVTSWGLG 402
Qy 181 CARPNKPGVYVRVSRFVTWIEGWENN 207
Db 403 CANAKPGVYVRVSRFVTWIEGWENN 429

RESULT 7
Q9QY29 PRELIMINARY; PRT; 310 AA.
AC Q9QY29;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Distal intestinal serine protease.
GN DISP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP MEDLINE=20246299; PubMed=10786627;
RA Shaw-Smith C.J., Coffey A.J., Leversha M., Freeman T.C., Bentley D.R.,
RA Walters J.R.;
RT "Characterization of a novel murine intestinal serine protease,
RT DISP.";
RL Biochim. Biophys. Acta 1490:131-136(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AJ243866; CAB56465.1; -.

```

DR HSP; P00763; IDPO.
 DR MEROPS; S01.042; --.
 DR MGD; MGI:1353645; Disp.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004235; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS0134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 310 AA; 33701 MW; P828EC7F6D25303F CRC64;

Query Match 41.7%; Score 464.5; DB 11; Length 310;
 Best Local Similarity 41.8%; Pred. No. 3.3e-39;
 Matches 92; Conservative 35; Mismatches 78; Indels 15; Gaps 3;

Qy 2 HFCGGTLLSPWLTAAHCLKSPSSYKVLGAHQEVNLEPHVQIEVSRLPLEPT-- 59
 Db 61 HICGSLIHVEVWLTAAHCFRRSLNPSFYHKVGGTLTSLLEPHSTLVAVRNIFVHPTYL 120
 Qy 60 -----RKDIALLLSSPAVITDKVIPACLPSPNVTWADRTCEFTGWTGCTGFGAGLLK 114
 Db 121 WADASSGDIALVQDTP-LRPSQFTPVCLPAQFTPLPGTVCWTGATQERDMSVLQ 179
 Qy 115 EAQLPVNIENKVNRYEFLNGR-----VQSTELCAGHLAGTDSQQSDSGGLVCFPEKD 167
 Db 180 ELAVPLDSEDCERKMYHTQGSLSGERIIQSDMLCAGYVEGQKDSQSDSGGLVCSINS 239
 Qy 168 KYILQGVTSNGLCARPKGVYVRSFVTVWIEGVNRN 207
 Db 240 SWTQVGITSMGICARPGVYTRVPTVYDVIQRIILAE 279

RESULT 8
 Q91XC4
 ID Q91XC4 PRELIMINARY; PRT; 310 AA.
 AC Q91XC4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to distal intestinal serine protease.
 GN DISP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RA Strausberg R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; BC010970; ARL10970.1; --.
 DR EMBL; BC040348; AAR40348.1; --.
 DR HSP; P00761; IAN1.
 DR MEROPS; S01.042; --.
 DR MGD; MGI:1353645; Disp.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 310 AA; 33707 MW; FA126747DEAB0AB6 CRC64;

Query Match 41.7%; Score 464.5; DB 11; Length 310;
 Best Local Similarity 41.8%; Pred. No. 3.3e-39;
 Matches 92; Conservative 35; Mismatches 78; Indels 15; Gaps 3;

Qy 2 HFCGGTLLSPWLTAAHCLKSPSSYKVLGAHQEVNLEPHVQIEVSRLPLEPT-- 59
 Db 61 HICGSLIHVEVWLTAAHCFRRSLNPSFYHKVGGTLTSLLEPHSTLVAVRNIFVHPTYL 120
 Qy 60 -----RKDIALLLSSPAVITDKVIPACLPSPNVTWADRTCEFTGWTGCTGFGAGLLK 114
 Db 121 WADASSGDIALVQDTP-LRPSQFTPVCLPAQFTPLPGTVCWTGATQERDMSVLQ 179
 Qy 115 EAQLPVNIENKVNRYEFLNGR-----VQSTELCAGHLAGTDSQQSDSGGLVCFPEKD 167
 Db 180 ELAVPLDSEDCERKMYHTQGSLSGERIIQSDMLCAGYVEGQKDSQSDSGGLVCSINS 239
 Qy 168 KYILQGVTSNGLCARPKGVYVRSFVTVWIEGVNRN 207
 Db 240 SWTQVGITSMGICARPGVYTRVPTVYDVIQRIILAE 279

RESULT 9
 Q812A6
 ID Q812A6 PRELIMINARY; PRT; 453 AA.
 AC Q812A6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transmembrane proteinase tmprss3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rao N.V., Rao G.N., Hoidal J.R.;
 RT "Genomic Organization of Murine Transmembrane Proteinases.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF479687; AAC33581.1; --.
 DR GO; GO:0016020; C-membrane; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0005044; F:scavenger receptor activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR001190; Srcr_receptor.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00192; LDLA; 1.
 DR SMART; SM00202; SR; 1.
 DR PROSITE; PS01209; LDLRA_1; 1.
 DR PROSITE; PS00068; LDLRA_2; 1.
 DR PROSITE; PS50068; LDLRA_2; 1.
 DR PROSITE; PS50287; SRCR_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 SQ SEQUENCE 453 AA; 49505 MW; 1EE7ECD6CB3DD894 CRC64;

Query Match 40.0%; Score 445.5; DB 11; Length 453;
 Best Local Similarity 42.6%; Pred. No. 4.7e-37;
 Matches 87; Conservative 37; Mismatches 75; Indels 5; Gaps 3;

```
Qy 2 HFCGGTILSPDWLTAAHCKLEKSPRPSSYKVLG--AHQEVNLEPHVQIEVSRLEFLEPT 59
Db 240 HLCGSIITPLUWITAAHCVYDLVHPKSWTVQVGLSLMDSVPFSLHVEKIYHSKYRKP 299
Qy 60 R-KDIALKLSPPAVITDKVIPACLPSPNYYVADRTECFITGWGETQ-GTFGAGLLKEA 116
Db 300 RLGNIDIALMKLSEPLTFDETQICLPNSEENFPDGKLCWTSWGATEDCGDASPVLNHA 359
Qy 117 QLPVNIENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQSGSGGLPVCFEKDKYILQGVTS 176
Db 360 AVPLISNKCINHRDVGIIISPSMLCAGYLKGGVDSCQSGSGGLPVCQERRLMLKVGATS 419
Qy 177 WGLGCAEVPKPGVYVRSRFTWI 200
Db 420 FGICAEVKNKPGVYTRITSLFDWI 443

RESULT 10
Q9DGR3 PRELIMINARY; PRT; 317 AA.
AC Q9DGR3; 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Embryonic serine protease-1.
GN XESP-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=20363741; PubMed=10903452;
RX Yamada K., Takabatake T., Takeshima K.;
RT "Isolation and characterization of three novel serine protease genes
from Xenopus laevis.";
RL Gene 252:209-216(2000);
DR EMBL; AB038496; BAB08216.1; -.
DR HSP; P00763; 1DPO.
DR MEROPS; S01.048; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 317 AA; 34413 MW; EEC78A9F46D138FE CRC64;

Query Match 39.5%; Score 440; DB 13; Length 317;
Best Local Similarity 40.5%; Pred. No. 1.1e-36;
Matches 87; Conservative 33; Mismatches 79; Indels 16; Gaps 3;

Qy 2 HFCGGTILSPDWLTAAHCKLEKSPRPSSYKVLG--AHQEVNLEPHVQIEVSRLEFLEPT--- 58
Db 64 HICGSIISDQWILATATIEHPDLPSGCGVRLGAYQLYKNPHEMTVKVDIIYINSEFN 123
Qy 59 ---TRKDIALKLSPPAVITDKVIPACLPSPNYYVADRTECFITGWGETQGTGFG---AGL 112
Db 124 GPGTSGDIALKLSPPKIFTEYIILPICLPASPVTFSSGTECWTGWGTSEVPLQYPAT 183
Qy 113 LKEAQLPVNIENKVCNRYEFLNG-----RVQSTELCAGHLAGGTDSCQSGSGGLPVCFE 165
Db 184 LQKVMVPIINRDSCEKMYHINSVISETEILLQSDQICAGYQAGKQDQCGDSCGGLPVCXI 243
```

```
Qy 166 KDKYILQGVTSWGLGCAEVPKPGVYVRSRFTWI 200
Db 244 QGFYQAGIVSWGERCAKRNKPGVYTFVPAYETWI 278

RESULT 11
Q80NM7 PRELIMINARY; PRT; 277 AA.
AC Q80NM7; 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Trypsin-6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Testis;
RT Wong G.W., Yasuda S., Li L., Stevens R.L.;
RT "Cloning and characterization of mouse trypsin-6 (mT6).";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY262280; AAP20885.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 277 AA; 29887 MW; 3310C067573E5E74 CRC64;

Query Match 39.3%; Score 437; DB 11; Length 277;
Best Local Similarity 39.4%; Pred. No. 1.9e-36;
Matches 86; Conservative 36; Mismatches 76; Indels 20; Gaps 4;

Qy 2 HFCGGTILSPDWLTAAHCKLEKSPRPSSYKVLG--AHQEVNLEPHVQIEVSRLEFLEPT--- 58
Db 57 HVCGSLIAPQWLVLTAGHCPRRWSEYSLVLCALSLDVRSSHELLVPLVRLVLLPDYS 116
Qy 59 ---TRKDIALKLSPPAVITDKVIPACLPSPNYYVADRTECFITGWGETQGTGAGL--- 112
Db 117 EDEARGDLALLQLRHVPVSLSTRIQPVCLPAPGSHPPGSPCWVTGWSLSP--GVPLPKG 174
Qy 113 --LKEAQLPVNIENKVCNRYEFLNGRVOSTE-----LCAGHLAGGTDSCQSGSGGLPVC 163
Db 175 RPLQGVRLDLSRACDRLYHGVANVPQGRIRVPLGNLCAGYRGRHKDQCGDSCGGLPVC 234
Qy 164 FEKDKYILQGVTSWGLGCAEVPKPGVYVRSRFTWI 201
Db 235 MESGHVVLVGVSWGKGCALPNRPGVYTNVAKYSPWIQ 272

RESULT 12
Q8NF86 PRELIMINARY; PRT; 284 AA.
AC Q8NF86; 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine protease EOS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```



```

RP SEQUENCE FROM N.A.
RA Darrow A.L., Qi J., Andrade-Gordon P., Chen C.;
RT "DNA encoding the human serine protease EOS."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF536382; AAN04055.1; -
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_SIA.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 284 AA; 30110 MW; FDF3F1750D569978 CRC64;

Query Match 39.3%; Score 437; DB 4; Length 284;
Best Local Similarity 38.8%; Pred. No. 1.9e-36;
Matches 87; Conservative 38; Mismatches 79; Indels 20; Gaps 4;

Qy 2 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRLEFLEP--- 58
Dy 60 HVCGSLIAPQWVLTAAHCFPRRLPAEYRVRLGALRLGTSPTLSVPVRRVLLPPDYS 119

Qy 59 ---TRKDIALLKSSPAVITDKVIPACLPSPNYVVDRTCEFTGWTGQTFGAGL--- 112
Dy 120 EDGARGDLALLQLRRPVLPSARQVCLPFGARPPPGTCRVTVGWSLRP--GVPLPEW 177

Qy 113 --LKEAOLPVLENKVCNRYEFLNGRVQSTE-----LCAGHLAGGTDSCGDSGGLVC 163
Dy 178 RPLQGVRRVLDLRTCDGLYHGVADVPQAEIRIVLPGLSCAGYPQHKDACQDSGGPLTC 237

Qy 164 FEKDKYILQGVTSWGLGCAKRNKPGVYVVRVSRFTWIEGVNRNN 207
Dy 238 LQSGSWLVGVVSGKGCALNRRGVTVSATVSPWIAQVTSN 281

RESULT 13
Q9BK47 PRELIMINARY; PRT; 267 AA.
AC Q9BK47;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sea STAR regeneration-associated protease SRAP.
OS Luidia foliolata.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC Asterozoa; Valvatacea; Paxillosida; Luidiidae; Luidia.
OX NCBI_TaxID=105861;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=2110042; PubMed=11179669;
RA Vickery M.C.L., McClintock J.B., Analer C.D.;
RT "Utilization of a novel deuterostome model for the study of
RT regeneration genetics: Molecular cloning of genes that are
RT differentially expressed during early stages of larval sea star
RT regeneration."
RL Gene 262:73-80(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF312826; AAK15274.1; -
DR HSP; P00763; 1DPO.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_SIA.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 267 AA; 28761 MW; 88F61A061921860C CRC64;

Query Match 39.0%; Score 434.5; DB 5; Length 267;
Best Local Similarity 39.2%; Pred. No. 3.2e-36;
Matches 83; Conservative 37; Mismatches 83; Indels 9; Gaps 3;

Qy 3 FCGGTLISPEWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRLEFLEP--- 59
Dy 58 FCGGTLISDEWAVSAACHFNHYGNINHYTAVVGAHRRDSVDSTQTTVGLGKVFVHESYDT 117

Qy 60 ---RKDIALLKSSPAVITDKVIPACLPSPNYVVDRTCEFTGWTGQTFGAGLLEA 116
Dy 118 STLDNDIALIKLSSPMSNYSVSCUPTA--ATPTGTECVTVGWGDOETAVDVDDPTLQOV 175

Qy 117 QLPVIENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCGDSGGLVC--FEKDKYILQGV 175
Dy 176 VPIISSEQCNRTWYGEINDNMICAGFEKGGKDCQDSGGPVCQSASGEYELGVV 235

Qy 176 SWGLGCAKRNKPGVYVVRVSRFTWIEGVNRNN 207
Dy 236 SWGYGCADARKPGYAKVLYNVSVNINLVARN 267

RESULT 14
Q8N171 PRELIMINARY; PRT; 327 AA.
AC Q8N171;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to protease, serine, 8 (Prostasin) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC036846; AAH36846.1; -
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_SIA.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease.
FT NON TER 1
SQ SEQUENCE 327 AA; 35106 MW; 60458ED817AC1CF3 CRC64;

Query Match 38.9%; Score 433; DB 4; Length 327;
Best Local Similarity 39.4%; Pred. No. 6e-36;
Matches 86; Conservative 36; Mismatches 76; Indels 20; Gaps 4;

Qy 2 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRLEFLEP--- 58
Dy 1 HFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRLEFLEP--- 58

```

Db 107 HVCGSLIAPQWVLTAAHCFRRALPAEYRVRGLALGLSTSPRTLSVPVRRVLLPDYS 166
 QY 59 ---TRKDIALKLSSPAVITDKVIPACLPSPNYVADRTECFITGCGTGTGAGL--- 112
 Db 167 EDGARGDALQLLRVPVLSARQVQCLPVGARPPGTFCRVTWGSLRP--GVPLPEW 224
 QY 113 --LKEAQLPVTIENKVCNRYEFLNGRVQSTE-----LCAGHLAGGTDSCQDGGSGGLVC 163
 Db 225 RPQGVVPLDSTCDGLYHVGADVPQAEIRIVLPGSLCAGYPOGHKDACQDGGSGGLTC 284
 QY 164 FEKDKYILQGVTSWGLGCARPNKPGVYVRSRFTWIE 201
 Db 285 LQSGSWLVGVSWGKGCALPNRPGVYTSVATYSPMIQ 322

RESULT 15

Q9XZM7
 ID Q9XZM7 PRELIMINARY; PRT; 581 AA.
 AC Q9XZM7;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cortical granule serine protease 1 precursor.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99303532; PubMed=10373300;
 RA Haley S.A., Wessel G.M.;
 RT "The cortical granule serine protease CGSP1 of the sea urchin,
 RT strongylocentrotus purpuratus, is autocatalytic and contains a low-
 RT density lipoprotein receptor-like domain.";
 RL Dev. Biol. 211:1-10(1999).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; AF149789; AAD37426.1; -;
 DR HSSP; P00760; IAQ7
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP1.
 DR Pfam; PF00057; ldl_recept_a; 4.
 DR Pfam; PF00089; trypsin_1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00192; LDLa; 5.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS01209; LDLRA_1; 3.
 DR PROSITE; PS00068; LDLRA_2; 1.
 DR PROSITE; PS02040; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR PROSITE; PS00092; TSP1; 1.
 KW Hydrolase; Protease; Serine protease; Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 581 CORTICAL GRANULE SERINE PROTEASE 1.
 SQ SEQUENCE 581 AA; 64438 MW; 692329218EBCA69F CRC64;

Query Match 38.9%; Score 432.5; DB 5; Length 581;
 Best Local Similarity 45.8%; Pred. No. 1.4e-35;
 Matches 99; Conservative 33; Mismatches 69; Indels 15; Gaps 8;

QY 4 CGGTLSIPWVLTAAHC-LEKSPRPSYKYVILGAHQEVNLEPH-----VQIEVSRLP 55
 Db 363 CGGTLDIPQVLTAAHCFMGPMMATSRWQVHLGKH-SVDFVPEAGSQHRLVREIFVHKCF 421
 QY 56 LE--PTRKDIALLKLSSPA-VITDKVIPACLPSPNYVADRTECFITGCGTGTGAG 111
 Db 422 GEHGGVGCIDIALILIDEPVPEQETGQINWACL-DEGMPLNDRTCYISGWSGVTEGNGGPD 480
 QY 112 LLKEAOLPVTIENKVCNRYEFLNGRVOSTELCAGHLAGGTDSCQDGGSGGLVCF-EKDKYI 170
 Db 481 VLHEARMPILPRICNYKKSNGKIEKTMLCAGHLEGGIDACQDGGSGGLSCLGPDHMY 540
 QY 171 LQGVTSWGLGCARPNKPGVYVRSRFTWIEGVNRN 206
 Db 541 VVGVTWGHGCAITANKPGVYTVKSSYLDWIDEMIH 576

Search completed: September 10, 2004, 14:55:06
 Job time : 129 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 13, 2004, 22:25:11 ; Search time 2502 Seconds
(without alignments)
2470.612 Million cell updates/sec

Title: US-09-992-095B-54
Perfect score: 1113
Sequence: 1 MHFCGGTLPSEWVLTAAHC.....GVYVRVSRFVTWIEGVNRNN 207

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_b/US0992095/runat_10092004_104838_4948/app.query.fasta_1.391
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0992095@cgn 1.1 3437 @runat_10092004_104838_4948 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rpd:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	1067	95.9	1201	9	AL579900	AL579900 AL579900
c 2	979	88.0	766	12	BI759134	BI759134 603042546
c 3	939	84.4	766	9	AV700701	AV700701 AV700701
c 4	851	76.5	523	14	CB164684	CB164684 K-EST0225
c 5	835	75.0	755	12	BI145879	BI145879 602911179
c 6	832	74.8	568	14	CD052213	CD052213 AB1346 ra
c 7	810.5	72.8	628	14	W98394	W98394 mg20b05.r1
c 8	798	71.7	654	9	AA032930	AA032930 mi22d06.r
c 9	785.5	70.6	772	14	CD297780	CD297780 AGENCOURT
c 10	784	70.4	711	14	CB948927	CB948927 AGENCOURT
c 11	782	70.3	564	10	BE665711	BE665711 154843 MA
c 12	779	70.0	504	14	N77239	N77239 yv44d01.r1
c 13	772	69.4	607	9	AV662061	AV662061 AV662061
c 14	768	69.0	622	9	AA268445	AA268445 va89g12.r
c 15	763	68.6	686	13	BA498968	BA498968 DXF2P779N
c 16	762	68.5	564	14	CB161593	CB161593 K-EST0221
c 17	757	68.0	727	14	CF178469	CF178469 807541 MA
c 18	747	67.1	1256	14	CD508756	CD508756 CDA92-F04
c 19	746	67.0	539	9	AA244769	AA244769 mx05g06.r
c 20	743	66.8	564	10	BE665645	BE665645 154756 MA
c 21	740	66.5	514	9	AA530434	AA530434 vj39c05.r
c 22	735.5	66.1	862	13	BA453606	BA453606 BX453606
c 23	735	66.0	785	12	BI554440	BI554440 603235751
c 24	733	65.9	622	9	AI061613	AI061613 HA0380 HU
c 25	725	65.1	749	14	CA376347	CA376347 654682 NC
c 26	714	64.2	762	12	BG428018	BG428018 602501424
c 27	714	64.2	763	13	BX870654	BX870654 BX870654
c 28	712	64.0	1312	14	CD495254	CD495254 CDA15-D07
c 29	709	63.7	1201	9	AL531542	AL531542 AL531542
c 30	705	63.4	751	13	BX883483	BX883483 BX883483
c 31	705	63.3	734	13	BX876806	BX876806 BX876806
c 32	701	63.0	1105	14	CF662376	CF662376 CGLU01a04
c 33	693	62.3	488	9	AA096737	AA096737 mo05g08.r
c 34	693	62.3	611	9	AV662084	AV662084 AV662084
c 35	689	61.9	463	14	W14157	W14157 mb19f08.r1
c 36	689	61.9	656	14	CB448635	CB448635 702784 MA
c 37	684	61.5	734	13	BX866810	BX866810 BX866810
c 38	681.5	61.2	433	14	H90220	H90220 yu89g08.r1
c 39	673	60.5	660	12	BM104371	BM104371 fv50f08.x
c 40	671.5	60.3	876	14	CB200754	CB200754 AGENCOURT
c 41	663	59.6	672	14	CF178792	CF178792 808309 MA
c 42	656.5	59.0	859	12	BG402207	BG402207 602465764
c 43	652.5	58.6	1180	14	CD508755	CD508755 CDA92-F04
c 44	652	58.6	676	9	AV172970	AV172970 AV172970
c 45	650	58.4	617	14	CB443102	CB443102 693965 MA

ALIGNMENTS

RESULT 1
AL579900/c
LOCUS AL579900 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CSDDJ002YJ21 3-PRIME, mRNA sequence.
ACCESSION AL579900
VERSION AL579900.2 GI:31318179
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)

AUTHORS TITLE JOURNAL COMMENT

Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12945394.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4076.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ002CE11NP1&cluster=4076.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ002CE11NP1.

FEATURES source

1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ002J21"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/clone_lib="JURKAT"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
Pred. No.: 1.85e-112 Length: 1201
Score: 1067.00 Matches: 206
Percent Similarity: 98.10% Conservatives: 0
Best Local Similarity: 98.10% Mismatches: 1
Query Match: 95.87% Indels: 3
DB: 9 Gaps: 0

US-09-992-095B-54 (1-207) x AL579900 (1-1201)

Qy 1 MethHisPheCysGlyGlyThrLeuIleSer-ProGluTrp-ValLeuThrAlaAlaHisC 20
Db 820 ATGCACCTCTGTGGAGGACCTTGATATCCCCAGAGTGGGTGTGCTGCTCCACT 761
Qy 20 YsLeuGluYsSerProArpSerSerTrpYsValIleLeuGlyAlaHisGlnGluV 40
Db 760 GCTTGGAGAGTCCCAAGGCCTTCATCCTACAGGTCATCTGGGTGCACCAAGAAG 701
Qy 40 al-AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThr 59
Db 700 TGGAAATCTCGAACCGCATGTCAGGAATAGAGTGTAGGCTGTCTTGGAGCCACA 641
Qy 60 ArgYsAspIleAlaLeuLeuYsLeuSerSerProAlaValIleThrAspYsValIle 79
Db 640 CGAAAAGATATGCTTGTCTAAGCTAAGCAGTCTCTCCGCTCATCTGCACAAAGTATC 581
Qy 80 ProAlaCysLeuProSerProAsnTrpYsValAlaAlaSerArgThrGluCysPheIleThr 99
Db 580 CCAGCTTGTCTGCCATCCCAAAATATGTGTGCTGACCGGACCGAATGTTTCATCACT 521
Qy 100 GlyTrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuYsGluAlaGlnLeuPro 119
Db 520 GGCTGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCT 461
Qy 120 ValIleGluAenLysValCysAsnArgTrpGluPheLeuAenGlyArgValGlnSerThr 139
Db 460 GTGAATTGAGATATAAGTGTGCATCGCTATGAGTTTCTGAATGGAGAGTCCATCCACC 401
Qy 140 GluLeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGly 159
Db 400 GAACTCTGTCTGGGCATTTGGCCGGAGGCACTGCACAGTTGCCAGGCTGACAGTGGAGGG 341

Qy 160 ProLeuValCysPheGluYsAspLysTrpIleLeuGlnGlyValThrSerTrpGlyLeu 179
Db 340 CCTCTGGTTTGCTTCGAGAGGACAATACATTTTACAGGAGTCACTTCTTGGGGTCTT 281
Qy 180 GlyCysAlaArgProAsnLysProGlyValTrpValArgValSerArgPheValThrTrp 199
Db 280 GGCTGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTTCAAGGTTTGTACTTGG 221
Qy 200 IleGluGlyValMetArgAenAen 207
Db 220 ATTGAGGAGTGTATGAGAAATAAT 197

RESULT 2

LOCUS BI759134 766 bp mRNA linear EST 25-SEP-2001
DEFINITION 603042546F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183003 5',
mRNA sequence.

ACCESSION BI759134

VERSION BI759134.1 GI:15750712

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 766)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11456 row: 1 column: 12

High quality sequence stop: 766.

FEATURES

source

1. 766

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5183003"

/lab_host="DH10B"

/clone_lib="NIH MGC 116"

/notes="Organ: pooled colon, kidney, stomach; Vector:

pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA

source anonymous pool of 3 colons, age 26 yo male, 49 yo

female, 71 yo male colon; 46 yo male kidney, and pool of 2

stomachs, 62 yo male and 70 yo female. Library is

oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.4 kb,

insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C.

Gruber (Invitrogen). Research Genetics tracking code

023. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 1.38e-102 Length: 766
Score: 979.00 Matches: 186
Percent Similarity: 99.47% Conservatives: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 87.96% Indels: 1
DB: 12 Gaps: 0

US-09-992-095B-54 (1-207) x BI759134 (1-766)

Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db |||||

```

Db      205 ATGCACCTCTCTGGAGGACCTTGTATATCCCAAGAGTGGGTGTGACTGTGCTGCCACTGC 264
Qy      21  LeuGluYsSerProArgProSerSerTyrylsValleLeuGlyAlaHisGlnGluVal 40
Db      265 TTGGAGAGTCCCCAAGGCTTCTATCTCTACAGGTCTCTCTGGGTGCACCAAGAGTG 324
Qy      41  AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60
Db      325 AATCTCGAACCGCATGTTTCAGGAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 384
Qy      61  LysSepileAlaLeuLeuYsLeuSerSerProAlaValleThrAspLysVallePro 80
Db      385 AAAGATATTCCTTGTCTAAAGCTAAGAGTCTCTCCGTCACTGACAAAGTAAATCCCA 444
Qy      81  AlaCysLeuProSerProAsnTyryValAlaAspArgThrGluCysPheIleThrGly 100
Db      445 GCTTGTCTGCCATCCCAATATGTCCTGCTGACCGACCGAATGTTTCATCACTGGC 504
Qy      101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuYsGluAlaGlnLeuProVal 120
Db      505 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 564
Qy      121 IleGluAsnLysValCysAsnArgTyryClnPheLeuGlnGlyArgValGlnSerThrGlu 140
Db      565 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCCAATCCACGAA 624
Qy      141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db      625 CTCTGTCTGGGCATTTGGCCGAGGCACTGACAGTTGGCCA-GGTGACAGTGGAGGGCT 683
Qy      161 LeuValCysPheGluYsAspLysTyryIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db      684 CTGGTTTCTTCGAGAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 743
Qy      181 CysAlaArgProAsnLysPro 187
Db      744 TGTGCACGCCCAATAGCCT 764

RESULT 3
AV700701/c 766 bp mRNA linear EST 16-JAN-2002
LOCUS AV700701 GK Homo sapiens cDNA clone GKAEH09 3', mRNA sequence.
DEFINITION AV700701
ACCESSION AV700701.1 GI:10302672
VERSION AV700701.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
XU, X., HUANG, J., XU, Z., QIAN, B., ZHU, Z., YAN, Q., CAI, T., ZHANG, X.,
XIAO, H., QU, J., LIU, P., HUANG, Q., CHENG, Z., LI, N., DU, J., HU, W.,
SHEN, K., LU, G., FU, G., ZHONG, M., XU, S., GU, W., HUANG, W., ZHAO, X.,
HU, G., GU, J., CHEN, Z. and HAN, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
MEDLINE 11752456
PUBMED
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhongjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1..766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
Location/Qualifiers
1..766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

/clone="GKAEH09"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKC"
/notes="vector: pbluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 6e-98 Length: 766
Score: 939.00 Matches: 177
Percent Similarity: 93.28% Conservative: 4
Best Local Similarity: 94.15% Mismatches: 7
Query Match: 84.37% Indels: 0
DB: 9 Gaps: 0

US-09-992-095B-54 (1-207) x AV700701 (1-766)
Qy      20  CysLeuGluYsSerProArgProSerSerTyrylsValleLeuGlyAlaHisGlnGlu 39
Db      764 TGCTTGAAGAAGTCCNCAGGCGCTTCTATCTCTCAAGGTCTATCTGGGTGCACCAAGAA 705
Qy      40  ValAsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThr 59
Db      704 GTGATTCGANC CGCATGTCAAGGAATAAGAGTGTCTAGGCTGTCTCTGAAGCCCA 645
Qy      60  ArgLysAspIleAlaLeuLeuYsLeuSerSerProAlaValleThrAspLysValle 79
Db      644 CGAAAGATATTCCTTGTCTTAAGTTAAGCAGTCTCTGGGTCACTGACAAAGTATTC 585
Qy      80  ProAlaCysLeuProSerProAsnTyryValAlaAspArgThrGluCysPheIleThr 99
Db      584 GCAGCTTGTCTGCCATCNCCAAATATGTCGTGCTGACCGGACCGAATGTTTCATCACT 525
Qy      100 GlyTrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuYsGluAlaGlnLeuPro 119
Db      524 GGCTGGGGAGAAACNCAGGTACTTTTGGAGCTGGCTTCTCAAGAAAGCCAGCTCCCT 465
Qy      120 ValIleGluAsnLysValCysAsnArgTyryGluPheLeuAsnGlyArgValGlnSerThr 139
Db      464 GTGATTCAGATTAAGTGTGCATCGCTATGAGTTTCTGAATGGAAGAGTCCATCCACC 405
Qy      140 GluLeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGly 159
Db      404 GAACCTCTGTCTGGGCATTTGGCCGAGGCACTGACAGTTGCCAGGTGACAGTGGAGGG 345
Qy      160 ProLeuValCysPheGluYsAspLysTyryIleLeuGlnGlyValThrSerTrpGlyLeu 179
Db      344 CCTCTGGTTTGTCTTCGAGAAGGACAAATACATTTTACAAGAGTCACTTCTTGGGGTCTT 285
Qy      180 GlyCysAlaArgProAsnLysProGlyValTyryValArgValSerArgPheValThrTrp 199
Db      284 GGCTGTGCACGCCCAATAGCTGGTGTCTATGTTCTGTGTTTCAGGTTTGTACTTGG 225
Qy      200 IleGluGlyValMetArgAsnAsn 207
Db      224 ATTGAGGGAGTGTGATGAGAAATAT 201

RESULT 4
CB164684 523 bp mRNA linear EST 30-JAN-2003
LOCUS CB164684
DEFINITION K-EST0225947 L17N670205n1 Homo sapiens cDNA clone
L17N670205n1-43-C07 5', mRNA sequence.
ACCESSION CB164684
VERSION CB164684.1 GI:28150810
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

```

Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)

TITLE JOURNAL COMMENT

Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 43 row: C column: 07
 High quality sequence stop: 523.

FEATURES

source

1. 523
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L17N670205n1-43-C07"
 /sex="P"
 /lab_host="Top10P"
 /clone_lib="L17N670205n1"

/note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
 Site 2: NotI; The library was contributed by the Soares
 laboratory and it was constructed as described by Bonaldo,
 M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
 6(9): 791-806. RNA was prepared from harvested cell
 culture."

ORIGIN

Alignment Scores:
 Pred. No.: 5,03e-88 Length: 523
 Score: 851.00 Matches: 161
 Percent Similarity: 98.77% Conservative: 0
 Best Local Similarity: 98.77% Mismatches: 0
 Query Match: 76.46% Indels: 2
 DB: 14 Gaps: 1

US-09-992-095B-54 (1-207) x CB164684 (1-523)

Qy 45 HisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArgLysAspIleAla 64
 Db 3 CATGTCAGGAATAGAAAGTCTCTAGG-----CTGGAGCCACACGAAAAGATATTGCC 56
 Qy 65 LeuLeuLysLeuSerProAlaValIleThrAspLysValIleProAlaCysLeuPro 84
 Db 57 TTGCTAAAGCTAAGCAGTCTCTGCCGTATCACTACACAAAGTAATCCAGCTTGTCTGCCA 116
 Qy 85 SerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGlyTrpGlyGluThr 104
 Db 117 TCCCCAAATTATGTGTGCTGACCGACCGAATGTTTCATCTGCTGGGGAGAAACC 176
 Qy 105 GlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProValIleGluAsnLys 124
 Db 177 CAAGGTACTTTTGGAGCTGGCTTCTCAAGGAGCCAGCCCTCTGTGATTGAGATAAA 236
 Qy 125 ValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeuCysAlaGly 144
 Db 237 GTGTGCAATCTATGAGTTTCTGAATGGAAGAGTCCCAATCCACCGAACTCTGTGTCTGG 296
 Qy 145 HisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPhe 164
 Db 297 CATTTGGCCGAGGACCTGACAGTTGCCAGGTGACAGTGGAGGCTCTGTGTTGCTTC 356
 Qy 165 GluLysAspTyrIleLeuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgPro 184
 Db 357 GAGAAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCC 416
 Qy 185 AsnLysProGlyValTyrValArgValSerArgPheValThrTrpIleGluGlyValMet 204
 Db 417 AATAAGCCTGTGTCTATGTTTCGTGTTTCAAGGTTTGTACTTGGATTGAGGGAGTGATG 476
 Qy 205 ArgAsnAsn 207

Db 477 AGAAATAAT 485

RESULT 5

B1145879

LOCUS

DEFINITION

602911179F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5052474 5',

mRNA sequence.

ACCESSION

B1145879

VERSION

B1145879.1 GI:14605880

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1. (Bases 1 to 755)

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

UNPUBLISHED (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM1142 row: m column: 19

High quality sequence stop: 734.

FEATURES

Location/Qualifiers

1..755

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5052474"

/lab_host="DH10B (TI phage-resistant)"

/clone_lib="NCI_CGAP_Li9"

/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.9 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.: 6.74e-86 Length: 755

Score: 835.00 Matches: 160

Percent Similarity: 87.11% Conservative: 9

Best Local Similarity: 82.47% Mismatches: 24

Query Match: 75.02% Indels: 1

DB: 12 Gaps: 0

US-09-992-095B-54 (1-207) x B1145879 (1-755)

Qy 15 LeuThrAlaAlaHisCysLeuGluLysSerProArgProSerSerTyrLysValIleLeu 34

Db 1 CTGACTGCTGCCACCTGTTGGAGAAATCTTCAAGACCTGAATCTTACAAGGTTATCTTG 60

Qy 35 GlyAlaHisGlnGluValAsnLeuGluProHisValGlnGlnGluValSerArgLeu 54

Db 61 GGTGGCGCAGCAAGATATATCCCTGGGTCCGATGTTCCAGGAATATCAGTAGCCAACTG 120

Qy 55 PheLeuGluProThrArgLysAspIleAlaLeuLeuLysLeuSerSerProAlaValIle 74

Db 121 ATCTTGGAGGCCAACACCGCTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Qy 75 ThrAspLysValIleProAlaCysLeuProSerProAsnTyrValValAlaAspArgThr 94

Db 181 ACCGATAAAGTCAATCCAGCTTGTCTGCCATCTCCAAATACATGTTGTCGCCGACA 240

Qy 95 GluCysPheIleThrGlyTrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLys 114

Db 241 ATATGTTACATCACCGGTGGGAGAGACTCAAGGAGACTTTTCGGTGGCGTCTCAAG 300
 Qy 115 GluAlaGlnLeuProValIleGluAenLysValCysAenArgTyrGluPheLeuAenGly 134
 Db 301 GAGGCTAGCTGCTGTGATTGAGNACAGGTGTGCACCGCGTGCAGTATCTGAACAC 360
 Qy 135 ArgValGlnSerThrGluLeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGln 154
 Db 361 AGAGTCAAAATCCACGAGCTCTGTGCGGGCAACTGGCTGGTGGCGTGCAGAGTCCAG 420
 Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174
 Db 421 GGCACAGTGGAGGACCTCTGGTTGCTTCGAGAGGCAAGTACATTTTCAAGGAGTC 480
 Qy 175 ThrSerTrpGlyLeuGlyCysAlaArgProAenLysProGlyValTyrValArgValSer 194
 Db 481 ACTTCTGGGGTCTTGGTGTGCTCGCCCAATAAGCTGGTGTCTAGTCTGTTCTCA 540
 Qy 195 ArgPheValThrTrpIleGluGlyValMetArgAenAen 207
 Db 541 CGGTTTGTGATTGACAAAGGAGATGAGGAATAAC 580

RESULT 6
 CD052213/c
 LOCUS
 DEFINITION Ab1346 rat regenerating liver after partial hepatectomy Rattus
 norvegicus cDNA, mRNA sequence.
 ACCESSION CD052213
 VERSION CD052213.1 GI:30534646
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 568)
 Xu, C.S., Li, Y.C., Li, W.Q., Lin, J.T., Zhang, H.Y., Rahman, S.,
 Zhang, J.B. and Wang, Q.N.
 A rat regenerating liver cDNA library
 Unpublished (2003)
 CONTACT: Cun-Shuan Xu, Yu-Chang Li, Wen-Qiang Li, Jun-Tang Lin,
 Hui-Yong Zhang, Salman Rahman, Jing-Bo Zhang, Qing-Nan Wang
 Henan Bioengineering Key Lab
 Henan Normal University
 Jianshe Road, Xinxiang City, P.R.China
 Tel: 00863733328084
 Fax: 00863733328524
 Email: xucs@263.net
 Seq primer: T7.

FEATURES
 source
 1..568
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /lab_host="JM109"
 /clone_lib="rat regenerating liver after partial
 hepatectomy"
 /note="vector: pGEM-T; A rat regenerating liver cDNA
 library was constructed after partial hepatectomy eight
 hours by SSH(Suppression subtractive hybridization). The
 genes are up-regulative in liver regeneration."

ORIGIN
 Alignment Scores:
 Pred. No.: 9.24e-86 Length: 568
 Score: 832.00 Matches: 159
 Percent Similarity: 88.36% Conservative: 8
 Best Local Similarity: 84.13% Mismatches: 22
 Query Match: 74.75% Indels: 0
 DB: 14 Gaps: 0

US-09-992-095B-54 (1-207) x CD052213 (1-568)
 Qy 7 ThrLeuIleSerProGluTyrTrpValLeuThrAlaAlaHisCysLeuGluLysSerProArg 26
 Db 568 ACTTATATATCCACAGAGTGGGTGCTGACTGCCGCTCAGTCTGGGAAATCTTCGAGA 509
 Qy 27 ProSerSerTyrLysValIleLeuGlyAlaHisGlnGluValAenLeuGluProHisVal 46
 Db 508 CCTGAATTTCTACAGGTTATCTCTGGAGCACACCAAGAACCAATCTTGGGTCTAGATGTT 449
 Qy 47 GlnGluIleGluValSerArgLeuPheLeuGluProThrArgLysAspIleAlaLeuLeu 66
 Db 448 CAGCAATAGCAGTAACCAAACTGCTTGGAAACCCACGACGCTGACATTTGCCCTGCTG 389
 Qy 67 LysLeuSerSerProAlaValIleThrAspLysValIleProAlaCysLeuProSerPro 86
 Db 388 AAGCTAAGCCGCCAGCCACCATCACAGATAATGTCATCCAGCTTGTCTGCCATCTCCA 329
 Qy 87 AsnTyrValValAlaAspArgThrGluCysPheIleThrGlyTyrGlyGluThrGlnGly 106
 Db 328 AATTATGTGGTGGCGACACACTGTGTATACATCCCGCTGGGGAGAACCAAGAGG 269
 Qy 107 ThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProValIleGluAenLysValCys 126
 Db 268 ACTCAGGTGCGGCCGCTCTCAAGAGGCCAGCTGCCGTGATCGAGAACAGGTGTGC 209
 Qy 127 AsnArgTyrGluPheLeuAenGlyArgValGlnSerThrGluLeuCysAlaGlyHisLeu 146
 Db 208 AACCGCGCTGATATCTAACAACAGAGTCAATCCACGAGCTCTGTCCGGGCATCTG 149
 Qy 147 AlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLys 166
 Db 148 GCTGGTGGCATCGACAGTTGCCAGGGCCACAGTGGAGGACCTCTGGTTGTTGTCGAGAAG 89
 Qy 167 AspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAenLys 186
 Db 88 GACAAGATATATTTTACAAGGAGTCACTCTTGGGGTCTTGGCTGTGTGCTGCCCAATG 29
 Qy 187 ProGlyValTyrValArgValSerArg 195
 Db 28 CTTGGTGTCTATGTTCTGTTTCCCGG 2

RESULT 7

W98394 628 bp mRNA linear EST 16-JUL-1996
 mg20b05.r1 Soares mouse embryo NM013.5 14.5 Mus musculus cDNA
 clone IMAGE:424305 5' similar to gb:J04766 Mouse plasminogen mRNA,
 complete cds (MOUSE);, mRNA sequence.

W98394
 W98394.1 GI:1428305
 EST.
 Mus musculus (house mouse)
 Mus musculus

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 628)

REFERENCE
 AUTHORS
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE
 JOURNAL
 COMMENT
 The WashU-HHMI Mouse EST Project
 Unpublished (1996)

CONTACT: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:258857

QY 26 ArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluValAsnLeuGluProHis 45
 DB 3 AGACCTGAATTCACAGGTTATCTCGGTGGCCAGCAAGAAATATATCCGTGGGTGGAT 62
 QY 46 ValGlnGluLeuGluValSerArgPheLeuGluProThrArgLysAspIleAlaLeu 65
 DB 63 GTTCAGGAATATCATAGTACCAACTGATCTTGGAGCCCAACACCGTGACATCCCTCG 122
 QY 66 LeuLysLeuSerSerProAlaValIleThrAspLysValIleProAlaCysLeuProSer 85
 DB 123 CTGAACCTAAGCGCGCCAGCCACCATCAGGATAAAGTCATTCACGCTTGTGCACTCT 182
 QY 86 ProAsnTyrValValAlaAspArgThrGluCysPheIleThrGlyTyrGlyGluThrGln 105
 DB 183 CCAATATCATGTGTGCTGACCGGACAATATGTTACATCACCCTGCGGTGGGAGAGACTCAA 242
 QY 106 GlyThrPheGlyAlaGlyLeuLysGluAlaGlnLeuProValIleGluAsnLysVal 125
 DB 243 GGGACTTTTCGGTGGCGGTGCTCTCAAGAGGCTCAGCTGCTGATTTGAGAACAGGTG 302
 QY 126 CysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeuCysAlaGlyHis 145
 DB 303 TGCAACCGGTGAGTATCTGACACAGAGTCAATCCAGGAGCTCTGTGCGGGCAA 362
 QY 146 LeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGlu 165
 DB 363 CTGGCTGGTGGCGTGCAGAGCTGCCAGGCGGACAGTGGAGGACCTCTGTTGCTTCGAG 422
 QY 166 LysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAsn 185
 DB 423 AAGGACAAAGTACATTTTACAGGAGTCACTCTTGGGGTCTTGGCTGTGCTCGCCCAAT 482
 QY 186 LysProGlyValTyrValArgValSerArgPheValThrTrpIleGluGlyValMetArg 205
 DB 483 AAGCTGGTGTCTACGTTCTGCTCAGCGTTTGTGATTTGATTTGAAGGAGATGAGG 542
 QY 206 AsnAsn 207
 DB 543 AATAAC 548

RESULT 9
 CD297780
 LOCUS
 DEFINITION AGNCOURT 14197274 NIH MGC 177 Mus musculus cDNA clone
 IMAGE:30381709 5', mRNA sequence.
 CD297780
 ACCESSION CD297780.1 GI:31077575
 VERSION EST.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Michael Brownstein
 cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: NDCM175 row: m column: 14
 High quality sequence stop: 425.
 Location/Qualifiers
 1. 772

FEATURES
 source

/organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30381709"
 /lab_host="DH10B (Ti-phage-resistant)"
 /clone_lib="NIH_MGC_177"
 /notes="Organ: liver; vector: pDNR-LIB; Site_1: Sfil
 (ggccattatggcc); Site_2: Sfil (ggcgcctggcc); cDNA made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTGATCAACGACAGTGGCCATTACGGCGGG-3' and
 5'-ATTCTAGAGCGGAGCGCGGACATG-dt(30)NN-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART kit and size-selected to contain the 0.5 kb
 size fraction. Library created in the laboratory of M.
 Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 3,85e-80 Length: 772
 Score: 785.50 Matches: 158
 Percent Similarity: 83.92% Conservat: 9
 Best Local Similarity: 79.40% Mismatches: 25
 Query Match: 70.58% Indels: 7
 DB: 14 Gaps: 2

US-09-992-095B-54 (1-207) x CD297780 (1-772)

QY 14 ValLeuThrAlaAlaHisCysLeuGluLysSerProArgProSerSerTyrLysValIle 33
 DB 5 GTTCTGACTGCTGCCACTGTTGGAGAAATCTTCAAGACCTGAATTTCTACAGGTTATC 64
 QY 34 LeuGlyAlaHisGlnGluValAsnLeuGluProHisValGlnGluIleGluLysSerArg 53
 DB 65 CTGGTGGCGGACGAGAGATATATCCGTGGGTGGATGTTTCAGGAAATATCATAGTACCA 124
 QY 54 LeuPheLeuGluProThrArgLysAspIleAlaLeuLysLysLeuSerSerProAlaVal 73
 DB 125 CTGATCTTGGAGCCCAACACCGTGACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 184
 QY 74 IleThrAspLysValIleProAlaCysLeuProSerProAsnTyrValValAlaAspArg 93
 DB 185 ATCAGGATTAAGTCATTCAGCTTGTGCTGCTTCTCAAAATTACATGTTGTGACCGG 244
 QY 94 ThrGluCysPheIleThrGlyTyrGlyThrGlnGlyThrPheGlyAlaGlyLeuLeu 113
 DB 245 ACAATATGTTACATCACCGCTGGGAGAGACTCAAGGAGCTTTTCGGTGGCGGTCTC 304
 QY 114 LysGluAlaGlnLeuProValIleGluAsnLysValCysAsnArgTyrGluPheLeuAsn 133
 DB 305 AAGGAGGCTCAGCTGCTGCTGATTGAGAACAAAGGTGTGCAACCGCTCGAGTATCTGAAC 364
 QY 134 GlyArgValGlnSerThrGluLeuCysAlaGlyHisLeuAlaGlyThrAspSerCys 153
 DB 365 AACAGAGTCAAAATCCAGGAGCTCTGTGCGGGCAACTGCTGCTGCTGCTGCTGCTGCTG 424
 QY 154 GlnGlyAspSerGlyGlyProLeuValCysPheGlyLysAspLysTyrIleLeuGlnGly 173
 DB 425 CAGGCGCAGCTGGAGGACCTCTGTTTGTTCGAGAGGAGCAAGTACATTTTACAGGA 484
 QY 174 ValThrSerTrpGlyLeuGlyCysAlaArg-ProAsnLysProGlyValTyrVal-ArgV 193
 DB 485 GTCACTTTTGGGTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
 QY 193 alserArgPheValThrTrp-----lleGluGlyValMetArgAsnAsn 207
 DB 545 TCTCACGGTTT-----TGGTGTGATTTGAAAGGAGAGATGGAGGAATAAC 591

RESULT 10

CE948927

LOCUS

DEFINITION

CE948927 711 bp mRNA linear EST 29-APR-2003
 AGENCOURT 13681788 NIH MGC 177 Mus musculus cDNA clone
 IMAGE:30307163 5', mRNA sequence.

```

ACCESSION CB948927
VERSION CB948927.1 GI:30201687
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 711)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Michael Brownstein
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM83 row: k column: 12
High quality sequence stop: 563.
Location/Qualifiers
1..711
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30307163"
/lab_host="DH10B (TI-phage-resistant)"
/notes="Organ: liver; Vector: pDNR-LIB; Site 1: Sfil
(ggccatagggc); Site 2: Sfil (ggccgctggc); cDNA made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAACGAGTGCCTACGCGCGG-3' and
5'-ATTTCAGGCGGCGGCGCATG-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4,99e-80 Length: 711
Score: 784.00 Matches: 148
Percent Similarity: 87.64% Conservatives: 8
Best Local Similarity: 83.15% Mismatches: 22
Query Match: 70.44% Indels: 0
DB: 14 Gaps: 0
US-09-992-095B-54 (1-207) x CB948927 (1-711)
Qy 30 TyrlvsVallleLeuGlyAlaHisGlnGluValAsnLeuGluProHisValGlnGluile 49
Db 9 TACAAGGTATCTCTGGTGGCGCACGAATATATCCGTGGTGGATGTTTCAGGAATA 68
Qy 50 GluValSerArgLeuPheLeuGluProThrArglyAspIleAlaLeuLeuLeuSer 69
Db 69 TCAGTAGCCAAATGATCTTGGAGCCCAACACCGTGACATTGGCCCTGCGTAAATAGC 128
Qy 70 SerProAlaValIleThrAspLysValIleProAlaCysLeuProSerProAsnTyVal 89
Db 129 CGCCAGTCACATCAGGATTAAGTCAATCCAGCTGTGTGCCATCTCCAAATTCATG 188
Qy 90 ValAlaSerArgThrGluCysPheIleThrGlyTrpGlyGluThrGlnGlyThPheGly 109
Db 189 GTTGCTGACCGGCAATATGTTATCATCCCGCTGGGAGAGACTCAAGGACTTCGGT 248
Qy 110 AlaGlyLeuLeuLysGluAlaGlnLeuProValIleGluAsnLysValCysAsnArgTy 129
Db 249 GCCGGTCGCTCTCAAGGAGGCTCAGCTGCTGTGATTGAGAACAAAGGGTGTCAACCGGTC 308
Qy 130 GluPheLeuAsnGlyArgValGlnSerThrGluLeuCysAlaGlyHisLeuAlaGlyGly 149

```

```

Db 309 GAGTATCTGAACAACAGAGTCAATCCAGGAGCTCTGTGCCGGCAACTGGTGTGGGC 368
Qy 150 ThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTy 169
Db 369 GTCGACAGCTGCCAGGCGGACAGTGGAGGACCTCTGTGTTGCTTCAGAGAGGACAGTAC 428
Qy 170 IleLeuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyVal 199
Db 429 ATTTTACAGGAGTCACTCTTGGGCTCTGGCTGTGCTCGCCCAATAAGCCTGGTGTGTC 488
Qy 190 TyrValArgValSerArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207
Db 489 TAGCTTCGTCTCACGGTTTGTGATTGGATTGAAAGGAGGATGAGGAATAAC 542
RESULT 11
BE665711
LOCUS 154843 MARC 4BOV Bos taurus cDNA 5', mRNA linear EST 25-APR-2001
DEFINITION BE665711
ACCESSION BE665711
VERSION BE665711.1 GI:10026302
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 564)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J. and Keele,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
11282978
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCCGAGTCACGACG
Plate: 68 row: D column: 20
Seq primer: ATTAGTGACACTATAG.
FEATURES
Location/Qualifiers
1..564
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4BOV"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 20 and day 40
embryos."
ORIGIN
Alignment Scores:
Pred. No.: 5.74e-80 Length: 564
Score: 782.00 Matches: 151
Percent Similarity: 86.56% Conservatives: 10
Best Local Similarity: 81.18% Mismatches: 25
Query Match: 70.26% Indels: 1
DB: 10 Gaps: 0

```

US-09-992-095B-54 (1-207) x BE665711 (1-564)

Qy 2 HisPheCysGlyThrLeuLeuSerProGluTrpValLeuThrAlaAlaHisCysLeu 21
 Db 7 CACTTCTGTGAGGAACCTCTATATCCCAAG-TGGGTGCTGACTGCTGCCATGCTGCTG 65
 Qy 22 GluYssSerProArgProSerSerTyriysValleuGlyAlaHisGlnGluValAsn 41
 Db 66 GACAACATTTTATAGCGCTGTCTATCTTACAAAGTCTCTCTGGTGCACACAGAGAAAGTC 125
 Qy 42 LeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArgLys 61
 Db 126 CGGAACAGAGTGTCCAGGAATACCGATGTCCAGGCTGTTCGGGAGCCCTCTCAGGCG 185
 Qy 62 AspIleAlaLeuLeuLysLeuSerSerProAlaValleuThrAspLysValleuProAla 81
 Db 186 GACATTGCTTACTCAAGCTGAGCAGACCTGCCATCATCACGAAGAGGTATATCCAGCT 245
 Qy 82 CysLeuProSerProAsnTyriysValAlaAlaAspArgThrGluCysPheLeuThrGlyTrp 101
 Db 246 TGTCTGCGCCACCCGGAACATACATGTTGACGCCGACAGATGTATCATCATCTGCTGCTG 305
 Qy 102 GlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProValle 121
 Db 306 GGAGAAACCCCAAGGCACCTTTTGGTGAGGGCTCTCTGAAGGAAGCGCACCTGCTGTGATC 365
 Qy 122 GluAsnLysValCysAsnArgTyriysGluPheLeuAsnGlyArgValGlnSerThrGluLeu 141
 Db 366 GAGNACAGGTGTGTATTCGACAGAGTATCTCACGGGAGGTCAAGCCACAGGAGCTG 425
 Qy 142 CysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeu 161
 Db 426 TGTGCGCGCATCTGATTGGAGGACTGACAGCTGCCAGGCTGACAGCGGGCGGCTCTG 485
 Qy 162 ValCysPheGluLysAspLysTyriysLeuGlnGlyValThrSerTrpGlyLeuGlyCys 181
 Db 486 GTCTGCTTTGAGAAGGCAATATACATCTGCAAGGAGTCACTTCTGGGGTCTTGGGCTG 545
 Qy 182 AlaArgProAsnLysPro 187
 Db 546 GCACGCCCAATAGCCT 563

RESULT 12

N77239
 LOCUS
 DEFINITION
 Yv44d01.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
 IMAGE:245569 5' similar to gb:X05199 PLASMINOGEN PRECURSOR
 (HUMAN); mRNA sequence.

N77239

Accession

Version

Keywords

Source

Organism

Homosapiens (human)

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

Homosapiens

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert length: 832 Std Error: 0.00
 Seq primer: reverse ET
 High quality sequence stop: 479.

FEATURES

source

1.504
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3794815"
 /db_xref="taxon:9606"
 /clone="IMAGE:245569"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares fetal liver spleen INFLS"
 /notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5' AACTGGAGAAATTAATAAGATCTTTTATTTTATTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Alignment Scores:
 Pred. No.: 1.06e-79 Length: 504
 Score: 779.00 Matches: 150
 Percent Similarity: 96.77% Conservative: 0
 Best Local Similarity: 96.77% Mismatches: 5
 Query Match: 69.99% Indels: 1
 DB: 14 Gaps: 0

US-09-992-095B-54 (1-207) x N77239 (1-504)

Qy 53 ArgLeuPheLeuGluProThrArgLysAspIleAlaLeuLeuLysLeuSerSerProAla 72
 Db 11 AGCTGTCTTGGAGCCACACGAAAGATATTGCTTGTCTAAGCTAAGCAGTCTGCTGCC 70
 Qy 73 ValileThrAspLysValleProAlaCysLeuProSerProAsnTyriysValleAlaAsp 92
 Db 71 GTCATCACTGACAAAGTAATCCAGCTGTCTGTCATCCCAATTTATGTGTGCTGCTGAC 130
 Qy 93 ArgThrGluCysPheLeuThrGlyTrpGlyGlnGlnGlyThrPheGlyAlaGlyLeu 112
 Db 131 CGGACCGAATGTTTCATCACTGCTGGGGGAAACCCCAAGTACTTTTGGAGCTGGCCTT 190
 Qy 113 LeuLysGluAlaGlnLeuProValleGluAsnLysValCysAsnArgTyriysGluPheLeu 132
 Db 191 CTCAGGAAGCCAGCTCCCTGTGATTGAGATAAAGTGTGCAATCGCTATGAGTTCCTG 250
 Qy 133 AsnGlyArgValGlnSerThrGluLeuCysAlaGlyHisLeuAlaGlyGlyThrAspSer 152
 Db 251 AATGGAAGAGTCCCAATCCACCGAATCTGTGCTGGGCAATTTGGCCGCGNC-CTGACACGT 309
 Qy 153 CysGlnGlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyriysLeuGln 172
 Db 310 TGCAGGGGTGACAGTGGAGTCTCTGCTGTTTGTCTTCAGGAAGGACAAATACATTTTACAA 369
 Qy 173 GlyValThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyriysValArg 192
 Db 370 GGAGTCACTTCTTGGGGTCTTGGCTGTGACGCCCAATTAAGCTGTGTGCTATGTTCTG 429
 Qy 193 ValSerArgPheValThrTrpIleGluGlyValMetArgAsn 207
 Db 430 GTTTCNAGGGTGTGTTACTTGGATTGAAGGAGTGTATGAGAAATAAT 474

RESULT 13

AV662061

LOCUS

DEFINITION

AV662061

607 bp

mRNA

linear

EST 16-JAN-2002

Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)

97044478

8889549

8889549

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

AV662061
 VERSION AV662061.1 GI:9883075
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 607)
 AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X., Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Gu, J., Chen, Z. and Han, Z.
 TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 MEDLINE 21625106
 PUBMED 11752456
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 source
 Location/Qualifiers
 1..607
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="G13AE06"
 /tissue_type="corresponding non cancerous liver tissue"
 /dev_stage="adult"
 /lab_host="SOLR"
 /clone_lib="GLC"
 /note="Vector: pbluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
 ORIGIN
 Alignment Scores:
 Pred. No.: 948-79 Length: 607
 Score: 772.00 Matches: 152
 Percent Similarity: 81.31% Conservative: 9
 Best Local Similarity: 76.77% Mismatches: 19
 Query Match: 69.36% Indels: 19
 DB: 9 Gaps: 1
 US-09-992-095B-54 (1-207) x AV662061 (1-607)
 Qy 1 MethisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 69 ATGCACCTCTGTGGAGGACCTTCATATGCCAGAGTGGGTGTGACTGTGCCCACTGC 128
 Qy 21 LeuGluIysSerProArgProSerSerTyrTyrValIleLeuGlyAlaHisGlnGluVal 40
 Db 129 TTGGAGAGTCCCAAGGCCTTCATCTACAGGTCTCTGGGTGCACACCAAGAAGTG 188
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 189 AATCTCGAACCGCATGTTTCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 248
 Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspIysValIlePro 80
 Db 249 -----ATCCCA 254
 Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 255 GCTTGTCTGCATCCCAATATGTGGTCCCGACCGACCGATATGTTTCATCCTGGC 314
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuIysGluAlaGlnLeuProVal 120
 Db 315 TGGGGAGAAACCAAGGTACTTTTGAAGTGCCTTTTAAGGAAGCCAGCTTCCTGTG 374

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 375 ATTGAGATTAAGTGTGCATTCCTATGAGTTCCTTGATGAGAGAGTCCCAATTCACCGAA 434
 Qy 141 LeuCyAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 435 CTCTGTCTGGCATTTGCCCGAAGGACTGACAAATTCGACGAGGTGACAGTGAAGGCT 494
 Qy 161 LeuValCysPheGluIysAspIysTyrIleLeuGlnGlyValThrSerThrGlyLeuGly 180
 Db 495 CTGGTTGTTTTCGGCAGGACAAATTCCTTTCCAGGGAGGAACCTTTTGGGTCCTGGC 554
 Qy 181 CyAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThr 198
 Db 555 TGAGCCACCCCAATTAACCTGCCT-TTGTACGGTTTCAAGATTTTAAACN 607
 RESULT 14
 LOCUS AA268445
 DEFINITION v998g12.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:747526 5', similar to gb:J04766 Mouse plasminogen mRNA, complete cds (MOUSE), mRNA sequence.
 ACCESSION AA268445
 VERSION AA268445.1 GI:1905181
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 622)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:456510
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 477.
 FEATURES
 source
 Location/Qualifiers
 1..622
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:747526"
 /tissue_type="Liver"
 /lab_host="DH108"
 /clone_lib="Soares mouse NML"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGGAGTAATCTTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,85e-78 Length: 622
 Score: 768.00 Matches: 150

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 13, 2004, 18:32:45 ; Search time 3729 Seconds
(without alignments)
2406.009 Million cell updates/sec

Title: US-09-992-095B-54
Perfect score: 1113
Sequence: 1 MHFCGGTLISPEWLTAAHC.....GVYVRVSRFVTWIEGVNRNN 207

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPRO_spool_h/US9992095/runat_10092004_104838_4938/app_query.fasta_1.391
-DB=GenEmbl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=text -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09992095 @CGN 1 1 3731 @runat_10092004_104838_4938 -NCPU=6 -ICPU=3
-NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_av:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_nu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1113	100.0	744	6	AX815991 Sequence
2	1113	100.0	750	6	AX463624 Sequence
3	1113	100.0	750	6	AX815990 Sequence
4	1113	100.0	783	6	AX815989 Sequence
5	1113	100.0	786	6	AX815988 Sequence
6	1113	100.0	999	6	AX815985 Sequence
7	1113	100.0	1005	6	AX815983 Sequence
8	1113	100.0	1011	6	AX815986 Sequence
9	1113	100.0	1017	6	AX815984 Sequence
10	1113	100.0	1038	6	AX815981 Sequence
11	1113	100.0	1041	6	AX815979 Sequence
12	1113	100.0	1047	6	AX463626 Sequence
13	1113	100.0	1047	6	AX815987 Sequence
14	1113	100.0	1050	6	AX815982 Sequence
15	1113	100.0	1053	6	AX815980 Sequence
16	1113	100.0	1302	6	AX815977 Sequence
17	1113	100.0	1314	6	AX815978 Sequence
18	1113	100.0	1724	6	AR363268 Sequence
19	1113	100.0	1784	9	HUMPMGM
20	1113	100.0	1907	6	AX616369 Sequence
21	1113	100.0	2145	6	AX815993 Sequence
22	1113	100.0	2296	6	AR105749 Sequence
23	1113	100.0	2296	6	I14104 Sequence 18
24	1113	100.0	2376	6	AX815992 Sequence
25	1113	100.0	2400	6	AX815934 Sequence
26	1113	100.0	2412	6	AX815936 Sequence
27	1113	100.0	2430	6	AR243324 Sequence
28	1113	100.0	2433	6	AX463622 Sequence
29	1113	100.0	2433	6	AX815938 Sequence
30	1113	100.0	2497	6	AX448883 Sequence
31	1113	100.0	2631	6	AX815940 Sequence
32	1113	100.0	2643	6	AX815942 Sequence
33	1113	100.0	2679	6	AR363269 Sequence
34	1113	100.0	2732	6	AX815951 Sequence
35	1113	100.0	2732	6	BD081408 Fused pro
36	1113	100.0	2732	9	HSPMGR
37	1113	100.0	2753	6	I45623 Sequence 1
38	1112	99.9	2497	6	AR037325 Sequence
39	1112	99.9	2497	6	AR082437 Sequence
40	1112	99.9	2497	6	AR085163 Sequence
41	1112	99.9	2497	6	BD105836 Novel ant
42	1112	99.9	2497	9	HUMELASM
43	1106	99.4	690	6	A37078 Sequence 1
44	1106	99.4	690	6	I51668 Sequence 1
45	1106	99.4	2730	9	BC060513 Homo sapi

ALIGNMENTS

AX815991
LOCUS AX815991 744 bp DNA linear PAT 09-DEC-2003
DEFINITION Sequence 64 from Patent WO03066842.
ACCESSION AX815991
VERSION AX815991.1 GI:39646630
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Susilo, R., Korting, H.C., Gassen, H.G., Hils, M. and Pasternack, R.
TITLE Method for producing recombinant proteins in micro-organisms
JOURNAL Patent: WO 03066842-A 64 14-AUG-2003;
Trommsdorff GmbH & Co. KG Arzneimittel (DE)
FEATURES
source Location/Qualifiers
1..744
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.67e-108 Length: 744
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-992-095B-54 (1-207) x AX815991 (1-744)
Qy 1 MethHiePheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 121 ATGCACCTTCGTGGAGGACCTTCATCTACAGGTCTATCCAGTGGGTGTGACTGCTGCCACTGC 180
Qy 21 LeuGluIysSerProArgProSerSerTyrIysValIleLeuGlyAlaHisGlnGluVal 40
Db 181 TTGGAGAAGTCCCAAGGCTTCATCTACAGGTCTATCCAGTGGGTGTGACTGCTGCCACTGC 240
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 241 AATCTCGAACCGCATGTTCCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 300
Qy 61 LysAspIleAlaLeuLeuIleLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 301 AAAGATATGCTTCCTGCTAAAGCTAAGCAGTCTCTCCGCTCATCTGACCAAGTAAATCCCA 360
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 361 GCTTGTCTGCCATCCCAATATGTTGGTCTGCTGACCGGACCGAATGTTTCATCTGCTGC 420
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuIysGluAlaGlnLeuProVal 120
Db 421 TGGGAGAAACCCCAAGTACTTTTGGAGTGGCTCTCAAGGAAGCCAGCTCCCTGTG 480
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 481 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCAATCCACCGAA 540
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 541 CTCTGTCTGGGCATTTGGCCGGAGGCACTGACAGTTGCCAGGCTGACAGTGGAGTCTCT 600
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 601 CTGCTTTGCTTCGAGAAGGACAAATACATTTTACAGAGTCACTTCTTGGGGTCTTGGC 660
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 661 TGTGACGCCCAATAAGCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 720
Qy 201 GluGlyValMetArgAsnAsn 207

|||||
721 GAGGAGTGTGAGAAATAAT 741
RESULT 2
LOCUS AX463624 750 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 3 from Patent WO0250290.
ACCESSION AX463624
VERSION AX463624.1 GI:21886384
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nagai, N., Laroche, Y. and Collen, D.J.
TITLE A yeast expression vector and a method of making a recombinant protein by expression in a yeast cell
JOURNAL Patent: WO 0250290-A 3 27-JUN-2002;
THROMB X NV (BE)
FEATURES
source Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
1..750
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD42144.1"
/db_xref="GI:21886385"
/db_xref="REMBL:CAD42144"
/translations="AFSDCKQKQVBPCKPCGRVGGCVAPHSWPMQVSLTRFGNH
FCGTLISPEWLTAAHCLKSPSSYKILGAHQVNLPHVQIEVEVSRLLPLEPTR
KDALLKLSPAVITDKVIPACLPSPNYVADRTGCTGWTGCTGCTGAGLKEAL
FVLENKVRNRYEFLNGVSTELCAGHLAGTSCQDSDSGVPLVCFKDKYILQGVTS
WGLGCRPNKPGVYVRSFVTWIEGVMNN"
ORIGIN
Alignment Scores:
Pred. No.: 6.73e-108 Length: 750
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-992-095B-54 (1-207) x AX463624 (1-750)
Qy 1 MethHiePheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 127 ATGCACCTTCGTGGAGGACCTTCATCTACAGTGGGTGTGACTGCTGCCACTGC 186
Qy 21 LeuGluIysSerProArgProSerSerTyrIysValIleLeuGlyAlaHisGlnGluVal 40
Db 187 TTGGAGAAGTCCCAAGGCTTCATCTCAAGTCTATCTCCGCTCATCTGAGTGGTGCACCAAGAAGTG 246
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 247 AATCTCGAACCGCATGTTCCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 306
Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 307 AAAGATATGCTTCCTGCTAAAGCTAAGCAGTCTCTCCGCTCATCTGACCAAGTAAATCCCA 366
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 367 GCTTGTCTGCCATCCCAATATGTTGGTCTGCTGACCGGACCGAATGTTTCATCTGCTGC 426
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuIysGluAlaGlnLeuProVal 120
Db 427 TGGGAGAAACCCCAAGTACTTTTGGAGTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 486
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140


```

Db      487 ATTGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCACCAGAA 546
Qy      141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db      547 CTCTGTGCTGGGCATTTGGCCGGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCT 606
Qy      161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db      607 CTGGTTTCTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 666
Qy      181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db      667 TGTGACGCCCAATAAGCTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 726
Qy      201 GluGlyValMetArgAsnAsn 207
Db      727 GAGGAGTGATGAGAAATAAT 747

RESULT 3
AX815990          750 bp      DNA      linear      PAT 09-DEC-2003
LOCUS
DEFINITION      Sequence 63 from Patent WO03066842.
ACCESSION      AX815990
VERSION        AX815990.1 GI:39646629
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS      Susilo, R., Kortling, H.C., Gassen, H.G., Hils, M. and Pasternack, R.
TITLE        Method for producing recombinant proteins in micro-organisms
JOURNAL      Patent: WO 03066842-A 63 14-AUG-2003;
Trommsdorff GmbH & Co. KG Arzneimittel (DE)
FEATURES
source
Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      6.73e-108      Length:      750
Score:          1113.00      Matches:      207
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6      Gaps:        0

US-09-992-095B-54 (1-207) x AX815990 (1-750)
Qy      1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db      127 ATGCACCTCTGTGAGGACCTTGATATCCCAAGTGGGTGTGACTGTGCCCACTGC 186
Qy      21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db      187 TTGGAGAAGTCCCAAGGCTTCTACCAAGGTCTATCTGGGTGCACACCAAGAGTG 246
Qy      41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db      247 AATCTCGAACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACAGA 306
Qy      61 LysAspIleAlaLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db      307 AAAGATATTGCTTGTCTAAAGCTAAGCAGTCTCCCGTCATCTGACAAAGTAAATCCCA 366
Qy      81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db      367 GCTTGTCTGCCATCTCCCAATATTGTGGTCTGCTGACCGGACCGAATGTTTCTACTGCGC 426
Qy      101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120

```

```

Db      427 TGGGGAGAAAACCCCAAGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCACCTCCCTGTG 486
Qy      121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db      487 ATTGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCACCAGAA 546
Qy      141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db      547 CTCTGTGCTGGGCATTTGGCCGGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCT 606
Qy      161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db      607 CTGGTTTCTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 666
Qy      181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db      667 TGTGACGCCCAATAAGCTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 726
Qy      201 GluGlyValMetArgAsnAsn 207
Db      727 GAGGAGTGATGAGAAATAAT 747

RESULT 4
AX815989          783 bp      DNA      linear      PAT 09-DEC-2003
LOCUS
DEFINITION      Sequence 62 from Patent WO03066842.
ACCESSION      AX815989
VERSION        AX815989.1 GI:39646628
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS      Susilo, R., Kortling, H.C., Gassen, H.G., Hils, M. and Pasternack, R.
TITLE        Method for producing recombinant proteins in micro-organisms
JOURNAL      Patent: WO 03066842-A 62 14-AUG-2003;
Trommsdorff GmbH & Co. KG Arzneimittel (DE)
FEATURES
source
Location/Qualifiers
1..783
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      7.08e-108      Length:      783
Score:          1113.00      Matches:      207
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6      Gaps:        0

US-09-992-095B-54 (1-207) x AX815989 (1-783)
Qy      1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db      160 ATGCACCTCTGTGAGGACCTTGATATCCCAAGTGGGTGTGACTGTGCCCACTGC 219
Qy      21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db      220 TTGGAGAAGTCCCAAGGCTTCTATCTCAAGGTCTATCTGGGTGCACACCAAGAGTG 279
Qy      41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db      280 AATCTCGAACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACAGA 339
Qy      61 LysAspIleAlaLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db      340 AAAGATATTGCTTGTCTAAAGCTAAGCAGTCTCCCGTCATCTGACAAAGTAAATCCCA 399
Qy      81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100

```

```

Db 400 GCTTGTCTGCCATCCCAATATATGTGTCGTGACCGGACCGAATGTTTCATCAGTGGC 459
Qy 101 TtpGlyThrGlnGlyThrPheGlyAlaGlyLeuLeuGlyAlaGlnLeuProVal 120
Db 460 TGGGGAGAAACCCCAAGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 519
Qy 121 IleGluAsnLysValCysAsnArgTyxGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 520 ATTGAGATAAAGTGTGCATCGTATGAGTTTCTGAATGGAGAGTCCATCCACCGAA 579
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 580 CTCTGTGCTGGGCAATTTGGCGGAGGCACACAGTTGCCAGGGTGACAGTGGAGTCTCT 639
Qy 161 LeuValCysPheGluLysAspLysTyxIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 640 CTGGTTTGGCTTCGAGAAGGCAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 699
Qy 181 CysAlaArgProAsnLysProGlyValTyxValArgValSerArgPheValThrTrpIle 200
Db 700 TGTGACGCGCCCAATAGCTGGTGTCTATGTTCTGTTTCAAGTTTGTACTTGGATT 759
Qy 201 GluGlyValMetArgAsnAsn 207
Db 760 GAGGAGTGTAGAAATAAT 780

RESULT 5
AX815988
LOCUS AX815988 786 bp DNA linear PAT 09-DEC-2003
DEFINITION Sequence 61 from Patent WO03066842.
ACCESSION AX815988
VERSION AX815988.1 GI:39646627
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Susilo,R., Kortling,H.C., Gassen,H.G., Hils,M. and Pasternack,R.
TITLE Method for producing recombinant proteins in micro-organisms
JOURNAL Patent: WO 03066842-A 61 14-AUG-2003;
Trommsdorff GmbH & Co.KG Arzneimittel (DE)
FEATURES
source 1..786
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 7.11e-108 Length: 786
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x AX815988 (1-786)
Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20
Db 163 ATGCACCTCTGTGGAGCACCTTGATATCCAGAGTGGGTGTGACTGTGCTGCCACTGC 222
Qy 21 LeuGluLysSerProArgProSerSerTyxIleLeuGlyAlaHisGlnGluVal 40
Db 223 TTGGAGAAGTCCCAAGCCCTTCACTCTACAAGTCACTCTCGGTGGTGACACCAAGAAGTG 282
Qy 41 AsnLeuGluProHisValGlnGluIleValSerArgLeuPheLeuGluProThrArg 60
Db 283 AATCTCGACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 342
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80

```

```

Db 343 AAAGATATTGCTTGTCTAAAGCTTAAGCAGTCTCTGCCGTCTATCTGACAAAGTAATCCCA 402
Qy 81 AlaCysLeuProSerProAsnArgTyxValValAlaAspArgThrGluCysPheIleThrGly 100
Db 403 GCTTGTCTGTCATCCCAATTAATGTGTCGTGACCGGACCGAATGTTTCATCAGTGGC 462
Qy 101 TtpGlyThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 463 TGGGGAGAAACCCCAAGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 522
Qy 121 IleGluAsnLysValCysAsnArgTyxGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 523 ATTGAGATAAAGTGTGCATCGTATGAGTTTCTGAATGGAGAGTCCATCCACCGAA 582
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 583 CTCTGTGCTGGGCAATTTGGCGGAGGCACACAGTTGCCAGGGTGACAGTGGAGTCTCT 642
Qy 161 LeuValCysPheGluLysAspLysTyxIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 643 CTGGTTTGGCTTCGAGAAGGCAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 702
Qy 181 CysAlaArgProAsnLysProGlyValTyxValArgValSerArgPheValThrTrpIle 200
Db 703 TGTGACGCGCCCAATAGCTGGTGTCTATGTTCTGTTTCAAGTTTGTACTTGGATT 762
Qy 201 GluGlyValMetArgAsnAsn 207
Db 763 GAGGAGTGTAGAAATAAT 783

RESULT 6
AX815985
LOCUS AX815985 999 bp DNA linear PAT 09-DEC-2003
DEFINITION Sequence 58 from Patent WO03066842.
ACCESSION AX815985
VERSION AX815985.1 GI:39646624
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Susilo,R., Kortling,H.C., Gassen,H.G., Hils,M. and Pasternack,R.
TITLE Method for producing recombinant proteins in micro-organisms
JOURNAL Patent: WO 03066842-A 58 14-AUG-2003;
Trommsdorff GmbH & Co.KG Arzneimittel (DE)
FEATURES
source 1..999
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.39e-108 Length: 999
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x AX815985 (1-999)
Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20
Db 376 ATGCACCTCTGTGGAGCACCTTGATATCCAGAGTGGGTGTGACTGTGCTGCCACTGC 435
Qy 21 LeuGluLysSerProArgProSerSerTyxIleLeuGlyAlaHisGlnGluVal 40
Db 436 TTGGAGAAGTCCCAAGCCCTTCACTCTACAAGTCACTCTCGGTGGTGACACCAAGAAGTG 495
Qy 41 AsnLeuGluProHisValGlnGluIleValSerArgLeuPheLeuGluProThrArg 60

```

```

Db      496 AATCTCGAACCGCATGTTTCAGGAATAGAAAGTGTCTAGGCTGTCTTGGAGCCACACGA 555
Qy      61 LysAspIleAlaLeuLeuLeuLeuSerProAlaValIleThrAspLysValIlePro 80
Db      556 AAGATATTGCTTGTCTTAAAGCTAAGCAGTCTGCGGTCTCATCTGACAAAGTAAATCCCA 615
Qy      81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db      616 GCTTGTCTGCCATCCCAAAATATATGTGCTGCTGACCGGACCGAATGTTTTCATCACTGGC 675
Qy      101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db      676 TGGGAGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAGGCCAGCTCCCTGTG 735
Qy      121 IleGluLeuLysValCysAsnArgTyrGluPheLeuLeuGlyArgValGlnSerThrGlu 140
Db      736 ATTGAGATTAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCATCCACCGAA 795
Qy      141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db      796 CTCGTGTCTGGGCATTTGGCGGAGGCACTGACAGTTGCCAGGCTGACAGTGGAGGTCTT 855
Qy      161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db      856 CTGTGTTTGTCTCGAAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 915
Qy      181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db      916 TGTGACGCCCAATAGCCGTGTCTATGTTCTGTGTTTCAGGTTTGTACTTGGATT 975
Qy      201 GluGlyValMetArgAsnAsn 207
Db      976 GAGGAGTGTAGAGAAATAAT 996

RESULT 7
AX815983
LOCUS      AX815983      1005 bp      DNA      linear      PAT 09-DEC-2003
DEFINITION Sequence 56 from Patent WO03066842.
ACCESSION AX815983
VERSION    AX815983.1 GI:39646622
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS    Susilo,R., Korting,H.C., Gassen,H.G., Hils,M. and Pasternack,R.
TITLE      Method for producing recombinant proteins in micro-organisms
JOURNAL    Patent: WO 03066842-A 56 14-AUG-2003;
Trommsdorff GmbH & Co.KG Arzneimittel (DE)
FEATURES   Location/Qualifiers
            source
            1..1005
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      9.46e-108      Length:      1005
Pred. No.:      1113.00      Matches:      207
Score:      100.00%      Conservative: 0
Percent Similarity: 100.00%      Mismatches: 0
Best Local Similarity: 100.00%      Indels: 0
Query Match:      100.00%      Gaps: 0
DB:      6

US-09-992-095B-54 (1-207) x AX815983 (1-1005)
Qy      1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db      382 ATGCACCTCTGTGAGGAGCACCTTGATATCCCAAGTGGGTGTTGACTGTGCTGCCATGCG 441
Qy      21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40

```

```

Db      442 TTGAGAAAGTCCCCAAGGCTTTCATCTCAAGGTCACTCTGGGTGCACCAAGAAGTG 501
Qy      41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db      502 AATCTCGAACCGCATGTTTCAGGAATAGAAAGTGTCTAGGCTGTCTTGGAGCCACACGA 561
Qy      61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db      562 AAGATATTGCTTGTCTTAAAGCTAAGCAGTCTGCGGTCTCATCTGACAAAGTAAATCCCA 621
Qy      81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db      622 GCTTGTCTGCCATCCCAAAATATATGTGCTGCTGACCGGACCGAATGTTTTCATCACTGGC 681
Qy      101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db      682 TGGGAGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAGGCCAGCTCCCTGTG 741
Qy      121 IleGluLeuLysValCysAsnArgTyrGluPheLeuLeuGlyArgValGlnSerThrGlu 140
Db      742 ATTGAGATTAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCATCCACCGAA 801
Qy      141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db      802 CTCGTGTCTGGGCATTTGGCGGAGGCACTGACAGTTGCCAGGCTGACAGTGGAGGTCTT 861
Qy      161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db      862 CTGTGTTTGTCTCGAAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 921
Qy      181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db      922 TGTGACGCCCAATAGCCGTGTCTATGTTCTGTGTTTCAGGTTTGTACTTGGATT 981
Qy      201 GluGlyValMetArgAsnAsn 207
Db      982 GAGGAGTGTAGAGAAATAAT 1002

RESULT 8
AX815986
LOCUS      AX815986      1011 bp      DNA      linear      PAT 09-DEC-2003
DEFINITION Sequence 59 from Patent WO03066842.
ACCESSION AX815986
VERSION    AX815986.1 GI:39646625
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS    Susilo,R., Korting,H.C., Gassen,H.G., Hils,M. and Pasternack,R.
TITLE      Method for producing recombinant proteins in micro-organisms
JOURNAL    Patent: WO 03066842-A 59 14-AUG-2003;
Trommsdorff GmbH & Co.KG Arzneimittel (DE)
FEATURES   Location/Qualifiers
            source
            1..1011
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      9.52e-108      Length:      1011
Pred. No.:      1113.00      Matches:      207
Score:      100.00%      Conservative: 0
Percent Similarity: 100.00%      Mismatches: 0
Best Local Similarity: 100.00%      Indels: 0
Query Match:      100.00%      Gaps: 0
DB:      6

US-09-992-095B-54 (1-207) x AX815986 (1-1011)
Qy      1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20

```

```

Db      388  ATGCACCTCTCTGGAGGACCTTGTATATCCAGAGTGGGTGTGACTGTCTGCCCACTGC 447
Qy      21  LeuGluLysSerProArgProSerSerTyrlsVallleLeuGlyAlaHisGlnGluVal 40
Db      448  TTGGAGAAGTCCCAAGCCCTTTCATCTCTACAAAGTCTATCTGGGTGCACCAAGAAAGTG 507
Qy      41  AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60
Db      508  AATCTGACACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACGCA 567
Qy      61  LysAspIleAlaLeuLeuLysLeuSerSerProAlaVallleThrAspLysValllePro 80
Db      568  AAAGATATTGCTTGTAAAGCTAAGCAGTCTCGCGTCATCACTGACAAAGTAATCCCA 627
Qy      81  AlaCysLeuProSerProAsnTyrlsValAlaAspArgThrGluCysPheIleThrGly 100
Db      628  GCTTGTCTGCCCATCCCAAAATTATGTGTCTGCTGACCGGACCGAATGTTTCATCACTGGC 687
Qy      101  TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db      688  TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCCACTCCCTGTG 747
Qy      121  IleGluAsnLysValCysAsnArgTyrlsPheLeuAsnGlyArgValGlnSerThrGlu 140
Db      748  ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 807
Qy      141  LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db      808  CTCTGTGTGGGCATTTGGCCGAGGACCTGACAGTTGCCAGGGTGACAGTGGAGGTCTCT 867
Qy      161  LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db      868  CTGGTTTCTTCGAGAAGGACAAATACATTTTACAGAGTCACTTCTTGGGTCCTTGGC 927
Qy      181  CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
Db      928  TGTGACGCCCAATAAGCCCTGGTGTCTATGTTCTGTGTTCAAGGTTTGTACTTGGATT 987
Qy      201  GluGlyValMetArgAsnAsn 207
Db      988  GAGGGAGTGTAGAAATAAT 1008

RESULT 9
AX815984
LOCUS      AX815984      1017 bp      DNA      linear      PAT 09-DEC-2003
DEFINITION Sequence 57 from Patent WO03066842.
ACCESSION AX815984
VERSION    AX815984.1 GI:39646623
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Susilo, R., Kortling, H.C., Gassen, H.G., Hils, M. and Pasternack, R.
TITLE      Method for producing recombinant proteins in micro-organisms
JOURNAL    Patent: WO 03066842-A 57 14-AUG-2003;
Trommsdorff GmbH & Co.KG Arzneimittel (DE)
FEATURES   source
            1. 1017
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      9.59e-108      Length:      1017
Pred. No.:      1113.00      Matches:      207
Score:      100.00%      Conservative: 0
Percent Similarity: 100.00%      Mismatches: 0
Best Local Similarity: 100.00%      Indels: 0
Query Match:      6      Gaps: 0
DB:

```

```

US-09-992-095B-54 (1-207) x AX815984 (1-1017)
Qy      1  MetHisPheCysGlyGlyThrLeuLysSerProGluTrpValleuThrAlaAlaHisCys 20
Db      394  ATGCACCTTCTGTGGAGGACCTTGTATATCCAGAGTGGGTGTGACTGTCTGCCCACTGC 453
Qy      21  LeuGluLysSerProArgProSerSerTyrlsVallleLeuGlyAlaHisGlnGluVal 40
Db      454  TTGGAGAAGTCCCAAGCCCTTTCATCTCTACAAAGTCTATCTGGGTGCACCAAGAAAGTG 513
Qy      41  AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60
Db      514  AATCTGCAACCGCATGTTTCAGGAAATAGAAAGTGTCTAGGCTGTCTTGGAGCCCAACGCA 573
Qy      61  LysAspIleAlaLeuLeuLysLeuSerSerProAlaVallleThrAspLysValllePro 80
Db      574  AAAGATATTGCTTGTCTTAAAGCTAAGCAGTCTCGCGTCATCACTGACAAAGTAATCCCA 633
Qy      81  AlaCysLeuProSerProAsnTyrlsValAlaAspArgThrGluCysPheIleThrGly 100
Db      634  GCTTGTCTGCCCATCCCAAAATTATGTGTCTGCTGACCGGACCGAATGTTTCATCACTGGC 693
Qy      101  TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db      694  TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCCACTCCCTGTG 753
Qy      121  IleGluAsnLysValCysAsnArgTyrlsPheLeuAsnGlyArgValGlnSerThrGlu 140
Db      754  ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 813
Qy      141  LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db      814  CTCTGTGTGGGCATTTGGCCGAGGACCTGACAGTTGCCAGGGTGACAGTGGAGGTCTCT 873
Qy      161  LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db      874  CTGGTTTCTTCGAGAAGGACAAATACATTTTACAGAGTCACTTCTTGGGTCCTTGGC 933
Qy      181  CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
Db      934  TGTGACGCCCAATAAGCCCTGGTGTCTATGTTCTGTGTTTCAAGGTTTGTACTTGGATT 993
Qy      201  GluGlyValMetArgAsnAsn 207
Db      994  GAGGGAGTGTAGAAATAAT 1014

RESULT 10
AX815981
LOCUS      AX815981      1038 bp      DNA      linear      PAT 09-DEC-2003
DEFINITION Sequence 54 from Patent WO03066842.
ACCESSION AX815981
VERSION    AX815981.1 GI:39646620
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Susilo, R., Kortling, H.C., Gassen, H.G., Hils, M. and Pasternack, R.
TITLE      Method for producing recombinant proteins in micro-organisms
JOURNAL    Patent: WO 03066842-A 54 14-AUG-2003;
Trommsdorff GmbH & Co.KG Arzneimittel (DE)
FEATURES   source
            1. 1038
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      9.82e-108      Length:      1038
Pred. No.:      1113.00      Matches:      207
Score:

```

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 6                             Gaps: 0

US-09-992-095B-54 (1-207) x AX815981 (1-1038)

Qy 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 415 ATGCACCTCTGTGGAGGACCTTGATATCCAGAGTGGGTGTGACTGTGCTGCCACTGC 474

Qy 21 LeuGluIleSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Db 475 TTGGAGAAGTCCCCCAAGCCCTTCACTCTACAGGTCTCTGGGTGCACCAAGAAGTG 534

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 535 AATCTCGAAGCCGATGTTCCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 594

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 595 AAGATATGCTTGTCTTAAAGCTAAGCAGTCTCTGCCGTCTCATCTGACAAAGTAATCCCA 654

Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 655 GCTTGTCTGCCATCCCCCAAAATTATGTGCTGCTGACCGGACCGAATGTTTCATCAGTGC 714

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 715 TGGGGAGAAACCCNAGGTACTTTTGGAGCTGGCTTCTCAAGNAGCCAGCTCCCTGTG 774

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 775 ATTGAGAATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 834

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 835 CTCTGTCTGGCATTTGGCGGAGGACACTGACAGTTGGCAGGGTGACGTGGAGGTCTCT 894

Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 895 CTGGTTGCTTCGAGAGGACAAATACATTTTACAGAGTCACTTCTTGGGTCTTGGC 954

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 955 TGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1014

Qy 201 GluGlyValMetArgAsnAsn 207
Db 1015 GAGGAGTGATGAGAAATAAT 1035

RESULT 11
LOCUS AX815979 1041 bp DNA linear PAT 09-DEC-2003
DEFINITION Sequence 52 from Patent WO03066842.
ACCESSION AX815979
VERSION AX815979.1 GI:39646618
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Susilo, R., Korting, H.C., Gassen, H.G., Hils, M. and Pasternack, R.
Method for producing recombinant proteins in micro-organisms
Patent: WO 03066842-A 52 14-AUG-2003;
Trommsdorff GmbH & Co. KG Arzneimittel (DE)
1. .1041
Location/Qualifiers
FEATURES
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 9.85e-108      Length: 1041
Pred. No.: 1113.00             Matches: 207
Score: 100.00%                 Conservative: 0
Percent Similarity: 100.00%     Mismatches: 0
Best Local Similarity: 100.00%  Indels: 0
Query Match: 100.00%           Gaps: 0
DB: 6

US-09-992-095B-54 (1-207) x AX815979 (1-1041)

Qy 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 418 ATGCACCTCTGTGGAGGACCTTGATATCCAGAGTGGGTGTGACTGTGCTGCCACTGC 477

Qy 21 LeuGluIleSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Db 478 TTGGAGAAGTCCCCCAAGCCCTTCACTCTACAGGTCTCTTGGGTGCACCAAGAAGTG 537

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 538 AATCTCGAAGCCGATGTTCCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 597

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 598 AAGATATGCTTGTCTTAAAGCTAAGCAGTCTCTGCCGTCTCATCTGACAAAGTAATCCCA 657

Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 658 GCTTGTCTGCCATCCCCCAAAATTATGTGCTGCTGACCGGACCGAATGTTTCATCAGTGC 717

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 718 TGGGGAGAAACCCNAGGTACTTTTGGAGCTGGCTTCTCAAGNAGCCAGCTCCCTGTG 777

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 778 ATTGAGAATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 837

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 838 CTCTGTCTGGCATTTGGCGGAGGACACTGACAGTTGCCAGGGTGACGTGGAGGTCTCT 897

Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 898 CTGGTTGCTTCGAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGTCTTGGC 957

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 958 TGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1017

Qy 201 GluGlyValMetArgAsnAsn 207
Db 1018 GAGGAGTGATGAGAAATAAT 1038

RESULT 12
LOCUS AX463626 1047 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 5 from Patent WO0250290.
ACCESSION AX463626
VERSION AX463626.1 GI:21886386
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Nagai, N., Laroche, Y. and Collen, D.J.
A yeast expression vector and a method of making a recombinant
protein by expression in a yeast cell
Patent: WO 0250290-A 5 27-JUN-2002;
THROMB X NV (BE)
Location/Qualifiers
FEATURES

```

```
source
1. .1047
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<1. .1047
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD42145.1"
/db_xref="GI:21886387"
/db_xref="REMRMBL:CAD42145"
/translacion="APPVLLPDPVETPSEEDCMFGKGYRGKRTVVTGTPCQDWA
AQEPRHSIFTPETNPGRVGVGCVAPHSWPQVSLRTFRGMHFGCGTGLIDSEWLTAAH
FDGCKPQVKPCPRVGVGCVAPHSWPQVSLRTFRGMHFGCGTGLIDSEWLTAAH
CLEKSPRSVSVTLGAHQENLEPHVQLEVSRLFLEPRKDIALKLSSPAVITDK
VIPACLSPNVVADRETCITGNETGTGAGLLKEAOLPVLENKVCNRYEFLNKR
VQSTELCAHLAGTDSQGDGSGPLVCFEKDKIILQGVTSWGLGCARPNKPGVYVRV
SRFVTWIEGVNRN"

ORIGIN
Alignment Scores:
Pred. No.: 9.92e-108 Length: 1047
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x AX463526 (1-1047)
Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 424 ATGCACCTTCTGTGGAGCCCTTGATATCCAGAGTGGGTGTGACTGTGCCACTGC 483
Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Db 484 TTGGAGAAGTCCCCAAGGCCCTTCATCTACAAGGTCTCTCGGTGGCACCAAGAAGTG 543
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 544 AATCTCGAACCGCATGTTTCAGGAATAGAAAGTCTAGGCTGTTCTTGGAGCCACACGA 603
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 604 AAAGATATTGCTTGTCTAAAGCTAAGCAGTCTGCGCTCATCTAGCAAGTAATCCCA 663
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 664 GCTTGTCTGCATCCCCCAAAATTATGTGCTGACCGGACCGAATGTTTCATCAGTGC 723
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyValAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 724 TGGGAGAAACCCAGGTACTTTTGGAGCTGGCCTTCTCAAGGAGCCAGTCCCTGTG 783
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 784 ATTGAGATAAAGTGTGCAATGCTAGTATCTGATGGAAGAGTCCATCCACCGAA 843
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 844 CTCTGTCTGGCATTTTGGCGGAGGACCTGACAGTTGGCAGGGTGACAGTGGAGGTCT 903
Qy 161 LeuValCysPheGluLysAspLysTyrlleuGlnGlyValThrSerTrpGlyLeuGly 180
Db 904 CTGTTTGTCTCGAAGAGCAATAATATTTCAGAGAGTCACTTCTTGGGGTCTTGGC 963
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 964 TGTGACGCCCAATAAGCTGTGTCTATGTTCTGTGTTTCAAGGTTTGTACTTGGATT 1023
Qy 201 GluGlyValMetArgAsnAsn 207
Db 1024 GAGGAGTGTATGAGAAATAAT 1044
```

```
RESULT 13
AX815987
LOCUS
DEFINITION
AX815987
Accession
AX815987
Version
AX815987.1
GI:39646626
Keywords
Homo sapiens (human)
Source
Homo sapiens
Organism
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Susilo,R., Korting,H.C., Gassen,H.G., Hils,M. and Pasternack,R.
AUTHORS
Method for producing recombinant proteins in micro-organisms
TITLE
Patent: WO 0306842-A 60 14-AUG-2003;
JOURNAL
Trommsdorff GmbH & Co.KG Arzneimittel (DE)
FEATURES
Location/Qualifiers
1. .1047
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.92e-108 Length: 1047
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x AX815987 (1-1047)
Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 424 ATGCACCTTCTGTGGAGCCCTTGATATCCAGAGTGGGTGTGACTGTGCCACTGC 483
Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Db 484 TTGGAGAAGTCCCCAAGGCCCTTCATCTACAAGGTCTCTCGGTGGCACCAAGAAGTG 543
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 544 AATCTCGAACCGCATGTTTCAGGAATAGAAAGTCTAGGCTGTTCTTGGAGCCACACGA 603
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 604 AAAGATATTGCTTGTCTAAAGCTAAGCAGTCTGCGCTCATCTAGCAAGTAATCCCA 663
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 664 GCTTGTCTGCATCCCCCAAAATTATGTGCTGACCGGACCGAATGTTTCATCAGTGC 723
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyValAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 724 TGGGAGAAACCCAGGTACTTTTGGAGCTGGCCTTCTCAAGGAGCCAGTCCCTGTG 783
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 784 ATTGAGATAAAGTGTGCAATGCTAGTATCTGATGGAAGAGTCCATCCACCGAA 843
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 844 CTCTGTCTGGCATTTTGGCGGAGGACCTGACAGTTGGCAGGGTGACAGTGGAGGTCT 903
Qy 161 LeuValCysPheGluLysAspLysTyrlleuGlnGlyValThrSerTrpGlyLeuGly 180
Db 904 CTGTTTGTCTCGAAGAGCAATAATATTTCAGAGAGTCACTTCTTGGGGTCTTGGC 963
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 964 TGTGACGCCCAATAAGCTGTGTCTATGTTCTGTGTTTCAAGGTTTGTACTTGGATT 1023
Qy 201 GluGlyValMetArgAsnAsn 207
Db 1024 GAGGAGTGTATGAGAAATAAT 1044
```

Qy 201 GluGlyValMetArgAsnAsn 207
 Db 1024 GAGGAGTGATGAGAAATAAT 1044

RESULT 14
 AX815982
 LOCUS AX815982 1050 bp DNA linear PAT 09-DEC-2003
 DEFINITION Sequence 55 from Patent WO03066842.
 ACCESSION AX815982
 VERSION AX815982.1 GI:39646621
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Suailo, R., Kortling, H.C., Gassen, H.G., Hills, M. and Pasternack, R.
 TITLE Method for producing recombinant proteins in micro-organisms
 JOURNAL Patent: WO 03066842-A 55 14-AUG-2003;
 Trommsdorff GmbH & Co.KG Arzneimittel (DE)

FEATURES
 source
 1. .1050
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 9.95e-108 Length: 1050
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x AX815982 (1-1050)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 427 ATGCACCTCTGTGGAGGACACCTTGATATCCACAGTGGGTGTGACTGTGCCACTGC 486

Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
 Db 487 TTGGAGAAGTCCCCAAGGCCCTTCATCTCTACAAGGTCTCTGGGTGCACACCAAGAAGTG 546

Qy 41 AsnLeuGluProHisValGlnGluIleValSerArgLeuPheLeuGluProThrArg 60
 Db 547 AATCTCGAACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 606

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 607 AAAGATATTGCTTAAAGCTAAGCAGTCTCGCGTCATCACTGACAAAGTAATCCCA 666

Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
 Db 667 GCTTGTCTGCCATCCCAATATATGTGCTGCTGACCGGACCGAATGTTTCATCACTGGC 726

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 727 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTCTCAAGGAAGCCAGCTCCCTGTG 786

Qy 121 IleGluAsnLysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 787 ATTGAGATAAAGTGTGCATCGTATGAGTTTCTGAATGGAAGAGTCCATCCACCGAA 846

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 847 CTCTGTCTGGCATTTGGCCGAGGACCTGACAGTTGCCAGGTGCAGTGGAGGTCTCT 906

Qy 161 LeuValCysPheGluLysAspLysTyrlsIleuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 907 CTGTGTTGCTTCGAGAAGGCAAAATACATTTTACAAAGAGTCACTTCTTGGGTCTTGGC 966

Qy 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
 Db 967 TGTGACGCCCCCAATAAGCCTGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1026

Qy 201 GluGlyValMetArgAsnAsn 207
 Db 1027 GAGGAGTGATGAGAAATAAT 1047

RESULT 15
 AX815980
 LOCUS AX815980 1053 bp DNA linear PAT 09-DEC-2003
 DEFINITION Sequence 53 from Patent WO03066842.
 ACCESSION AX815980
 VERSION AX815980.1 GI:39646619
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Suailo, R., Kortling, H.C., Gassen, H.G., Hills, M. and Pasternack, R.
 TITLE Method for producing recombinant proteins in micro-organisms
 JOURNAL Patent: WO 03066842-A 53 14-AUG-2003;
 Trommsdorff GmbH & Co.KG Arzneimittel (DE)

FEATURES
 source
 1. .1053
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 9.98e-108 Length: 1053
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x AX815980 (1-1053)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 430 ATGCACCTCTGTGGAGGACACCTTGATATCCACAGTGGGTGTGACTGTGCCACTGC 489

Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
 Db 490 TTGGAGAAGTCCCCAAGGCCCTTCATCTCTACAAGGTCTCTGGGTGCACACCAAGAAGTG 549

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 550 AATCTCGAACCGCATGTTTCAGGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 609

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 610 AAAGATATTGCTTAAAGCTAAGCAGTCTCGCGTCATCACTGACAAAGTAATCCCA 669

Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
 Db 670 GCTTGTCTGCCATCCCAATATATGTGCTGCTGACCGGACCGAATGTTTCTCACTGGC 729

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 730 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTCTCAAGGAAGCCAGCTCCCTGTG 789

Qy 121 IleGluAsnLysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 790 ATTGAGATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAAGAGTCCATCCACCGAA 849

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 850 CTCTGTCTGGCATTTGGCCGAGGACCTGACAGTTGCCAGGTGCAGTGGAGGTCTCT 909

Qy	161	LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly	180
Db	910	CTGTTTGCTTCGAGAGGACAAATACATTTTACAAGGAGTCACCTTCTTGGGGCTTGGC	969
Qy	181	CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle	200
Db	970	TGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT	1029
Qy	201	GluGlyValMetArgAsnAsn	207
Db	1030	GAGGAGTGATGAGAAATAAT	1050

Search completed: September 14, 2004, 00:35:15
Job time : 3736 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 13, 2004, 18:27:10 ; Search time 391 Seconds
(without alignments)
2249.046 Million cell updates/sec

Title: US-09-992-095B-54
Perfect score: 1113
Sequence: 1 MHFCGGLISPEWVIAAHC.....GVYRVSRFVTWIEGVNRNN 207

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO_spool_h/US0992095/runat_10092004_104837_4930/app_query.fasta_1.391
-DB=Geneseq_29Jan04 -QW=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pt0 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0992095 @CGN 1.1 470 @runat_10092004_104837_4930 -NCPU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1113	100.0	750	6 ABN89460	Abn89460 Human mic
2	1113	100.0	1047	6 ABN89461	Abn89461 Human min
3	1113	100.0	1724	2 AAQ40318	Aaq40318 Sequence
4	1113	100.0	1907	7 ACC51086	Acc51086 Human pla
5	1113	100.0	2296	2 AAQ40258	Aaq40258 Plasmid p
6	1113	100.0	2296	3 AAA89829	Aaa89829 Plasmid p
7	1113	100.0	2296	4 AAS12747	Aas12747 Plasmid p
8	1113	100.0	2430	7 ABX15877	Abx15877 DNA encod

9	1113	100.0	2433	2 AAX77711	Aax77711 Human pla
10	1113	100.0	2433	6 ABN89459	Abn89459 Human pla
11	1113	100.0	2679	2 AAQ40319	Aaq40319 Sequence
12	1113	100.0	2732	2 AAX35376	Aax35376 SEQ ID 51
13	1113	100.0	2732	6 ABN81696	Abn81696 Human pla
14	1113	100.0	2732	7 AAD56132	Aad56132 Human pla
15	1113	100.0	2732	7 ABX74464	Abx74464 Human CDN
16	1113	100.0	2732	7 ABA00770	Aba00770 Human pla
17	1113	100.0	6010	2 AAQ11998	Aaq11998 Human pla
18	1113	100.0	6020	2 AAQ06648	Aaq06648 Plasmid p
19	1112	99.9	2497	3 AAA52284	Aaa52284 DNA encod
20	1112	99.9	2497	6 ABN85296	Abn85296 Human pla
21	1112	99.9	2497	6 ABL54038	Ab154038 Human pla
22	1106	99.4	690	2 AAQ55506	Aaq55506 Sequence
23	1106	99.4	2753	2 AAQ12547	Aaq12547 Encodes P
24	1106	99.4	2756	2 AAQ12542	Aaq12542 Encodes P
25	1106	99.4	2756	2 AAQ12553	Aaq12553 Encodes P
26	1106	99.4	2756	2 AAQ12548	Aaq12548 Encodes P
27	1106	99.4	2756	2 AAQ12554	Aaq12554 Encodes P
28	1106	99.4	2756	2 AAQ12552	Aaq12552 Encodes P
29	1106	99.4	2759	2 AAQ12543	Aaq12543 Encodes P
30	1106	99.4	2759	2 AAQ12557	Aaq12557 Encodes P
31	1106	99.4	2759	2 AAQ12549	Aaq12549 Encodes P
32	1106	99.4	2759	2 AAQ12550	Aaq12550 Encodes P
33	1106	99.4	2759	2 AAQ12558	Aaq12558 Encodes P
34	1106	99.4	2762	2 AAQ12544	Aaq12544 Encodes P
35	1106	99.4	2762	2 AAQ12545	Aaq12545 Encodes P
36	1106	99.4	2762	2 AAQ12546	Aaq12546 Encodes P
37	1106	99.4	2765	2 AAQ12556	Aaq12556 Encodes P
38	1106	99.4	2771	2 AAQ12551	Aaq12551 Encodes P
39	1106	99.4	2771	2 AAQ12555	Aaq12555 Encodes P
40	1088	97.8	2679	2 AAT89686	Aat89686 Plasmid p
C 41	1006.5	90.4	2756	5 AAS80439	Aas80439 DNA encod
42	1006.5	90.4	2756	5 AAS66112	Aas66112 DNA encod
43	933.5	83.9	10422	7 ACA64741	Aca64741 Apolipop
44	933.5	83.9	13938	7 ACC47280	Acc47280 Human apo
45	925.5	83.2	13869	7 ACC49787	Acc49787 Human lip

ALIGNMENTS

RESULT 1
ABN89460
ID ABN89460 standard; cDNA; 750 BP.
XX AC ABN89460;
XX
DT 02-SEP-2002 (first entry)
XX
DE Human microplasmaenogen encoding cDNA SEQ ID NO:3.
XX
KW Human; plasmaenogen; microplasmaenogen; miniplasmaenogen; yeast; vector;
KW yeast expression vector; cardiant; thrombolytic; cerebroprotective;
KW fibrin proteolysis; kringle domain; thromboembolic disease;
KW focal cerebral ischaemic infarction; ischaemic stroke;
KW arterial thrombotic disease; peripheral arterial occlusive disease;
KW acute myocardial infarction; gene; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 1..750
FT /*tag= a
FT /partial
FT /product= "microplasmaenogen"
FT /note= "no start codon given"

XX WO200250290-A1.
XX 27-JUN-2002.
XX 20-DEC-2001; 2001WO-BE000217.
XX

PR 21-DEC-2000; 2000GB-00031196.
 PR 09-JUL-2001; 2001GB-00016690.
 PR 09-JUL-2001; 2001GB-00016702.
 XX (THRO-) THROMB-X NV.
 XX Collen DJ, Nagai N, Laroche Y;
 XX WPI; 2002-500632/53.
 DR P-PSDB; ABB81497.
 XX Novel expression vector for expressing mammalian plasminogen derivatives
 PT in yeast, has nucleotide sequence coding for catalytic domain of
 PT plasminogen and/or coding for kringle domains of plasminogen linked to
 PT promoter.
 XX Claim 7; Fig 3; 6lpp; English.
 XX The present invention describes a yeast expression vector (I) comprising
 CC a mammalian nucleotide sequence operably linked to a promoter, where the
 CC mammalian nucleotide sequence codes for the catalytic domain of
 CC plasminogen and further optionally codes for one or more kringle domains
 CC of plasminogen, its mutants or hybrids. (I) has cardiant, thrombolytic
 CC and cerebroprotective activities, and can mediate fibrin proteolysis.
 CC (I) can be used useful for treating a thromboembolic disease in a mammal.
 CC The mammalian protein expressed by (I) is useful for treating focal
 CC cerebral ischaemic infarction (ischaemic stroke) or arterial thrombotic
 CC diseases such as peripheral arterial occlusive disease or acute
 CC myocardial infarction. The present sequence encodes human
 CC microplasminogen, which is used in the exemplification of the present
 CC invention
 XX SQ Sequence 750 BP; 178 A; 173 C; 207 G; 192 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.06e-112 Length: 750
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-992-095b-54 (1-207) x ABB81496 (1-750)
 QY 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 DB 127 ATGCACCTTCTGTGGAGGACCTTGTATATCCAGAGTGGTGTGACTGTGCCCACTGC 186
 QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
 DB 187 TTGGAGAAGTCCCAAGGCTTCTATCTACAGGTCTCTGGTGTGCACACCAAGAGTG 246
 QY 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 DB 247 AATCTCGAACCGCATGTTCCAGGAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 306
 QY 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 DB 307 AAAGATATTCCTTGTCTAAAGCTAAGCAGTCTCTGCGCTCATCTGACAAAGTAATCCCA 366
 QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 DB 367 GCTTGTCTGCATCCCCCAATATGTGTGTGCTGACCGACCGAATGTTTTCATCTGCTGC 426
 QY 101 TrpGlyGluThrGlnGlyThrPheGlyValGlyLeuLeuLysGluAlaGlnLeuProVal 120
 DB 427 TGGGAGAACCCCAAGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 486
 QY 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 DB 487 ATTGAGAATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAAGAGTCCATCCACGAA 546
 QY 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160

Db 547 CTCTGTGCTGGCATTTGGCCGAGGACCTGACAGTTCAGGAGTGCAGTGGAGTCTCT 606
 QY 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTyrGlyLeuGly 180
 Db 607 CTGGTTTGTCTTCGAGAAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 666
 QY 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 Db 667 TGTGCAGGCCCAATAGCTGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 726
 QY 201 GluGlyValMetArgAsnAsn 207
 Db 727 GAGGGAGTGATGAGAAATAT 747
 RESULT 2
 ABB81496
 ID ABB81496 standard; cDNA; 1047 BP.
 XX AC ABB81496;
 XX 02-SEP-2002 (first entry)
 DE Human miniplasminogen encoding cDNA SEQ ID NO:5.
 XX Human; plasminogen; microplasminogen; miniplasminogen; yeast; vector;
 KW yeast expression vector; cardiant; thrombolytic; cerebroprotective;
 KW fibrin proteolysis; kringle domain; thromboembolic disease;
 KW focal cerebral ischaemic infarction; ischaemic stroke;
 KW arterial thrombotic disease; peripheral arterial occlusive disease;
 KW acute myocardial infarction; gene; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 1.1047
 CDS /*tag= a
 FT /partial
 FT /product= "miniplasminogen"
 FT /note= "no start codon given"
 XX WO200250290-A1.
 PD 27-JUN-2002.
 XX 20-DEC-2001; 2001WO-BE000217.
 XX 21-DEC-2000; 2000GB-00031196.
 PR 09-JUL-2001; 2001GB-00016690.
 PR 09-JUL-2001; 2001GB-00016702.
 XX (THRO-) THROMB-X NV.
 XX Collen DJ, Nagai N, Laroche Y;
 XX WPI; 2002-500632/53.
 DR P-PSDB; ABB81497.
 XX Novel expression vector for expressing mammalian plasminogen derivatives
 PT in yeast, has nucleotide sequence coding for catalytic domain of
 PT plasminogen and/or coding for kringle domains of plasminogen linked to
 PT promoter.
 XX Claim 8; Fig 4; 6lpp; English.
 XX The present invention describes a yeast expression vector (I) comprising
 CC a mammalian nucleotide sequence operably linked to a promoter, where the
 CC mammalian nucleotide sequence codes for the catalytic domain of
 CC plasminogen and further optionally codes for one or more kringle domains
 CC of plasminogen, its mutants or hybrids. (I) has cardiant, thrombolytic
 CC and cerebroprotective activities, and can mediate fibrin proteolysis.
 CC (I) can be used useful for treating a thromboembolic disease in a mammal.
 CC The mammalian protein expressed by (I) is useful for treating focal

CC cerebral ischaemic infarction (ischaemic stroke) or arterial thrombotic
 CC diseases such as peripheral arterial occlusive disease or acute
 CC myocardial infarction. The present sequence encodes human
 CC miniplasminogen, which is used in the exemplification of the present
 CC invention

SQ Sequence 1047 BP; 254 A; 250 C; 288 G; 255 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6.35e-112 Length: 1047
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x ABN89461 (1-1047)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 424 ATGCACCTTCTGTGGAGGACCTTGATATCCCGAGGTGGGTGTGACTGTGCCCACTGC 483
 Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
 Db 484 TTGGAGAGTCCCCCAAGCCCTTCATCTACAGGTCTATCCGGGTGCACCAAGAAGTG 543
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 544 AATCTCGNACCGCATGTTCCAGGAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 603
 Qy 61 LysAspIleAlaLeuLeuLeuLeuSerProAlaValIleThrAspLysValIlePro 80
 Db 604 AAAGATATGCTTGTCTAAAGCTAAGCAGTCTCCGCTCATCTACGACAAAGTAATCCCA 663
 Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAspArgThrGluCysPheIleThrGly 100
 Db 664 GCTTGTCTGCATCCCAATATGTGTCTGCTGACCGGACCGAATGTTTTCATCACTGGC 723
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 724 TGGGGAGAAACCCNAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGTCCCTGTG 783
 Qy 121 IleGluAsnLysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 784 ATTGAGATAAAGTGCATCGTATGAGTTTCTGAATGCAAGAGTCCAATCCACCGAA 843
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 844 CTCTGTCTGGGCATTTGGCCGAGGACGACAGTTGCCAGGCTGACAGTGGAGGTCTCT 903
 Qy 161 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 904 CTGGTTTGTCTCGAAGAGCAATACATTTACAGAGTCACTTCTTGGGTCTTGGC 963
 Qy 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
 Db 964 TGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1023
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db 1024 GAGGGAGTGTAGAGAAATAT 1044

RESULT 3

ID AAQ40318 standard; DNA; 1724 BP.

XX AC AAQ40318;

XX DT 17-AUG-1993 (first entry)

XX DE Sequence encoding tissue plasminogen activator (t-PA)/plasminogen hybrid
 DE protein.

XX

KW Zymogen; fibrinolytic activity; cleavage; ss.

OS Synthetic.

FH Key Location/Qualifiers
 FT CDS 16..1656
 FT /*tag= a

PN US5200340-A.

PD 06-APR-1993.

XX 22-MAY-1987; 87US-00053412.

XX 22-MAY-1987; 87US-00053412.

XX (ZYMO) ZYMOGENETICS INC.

XX Foster DC, Mulvihill ER, Ohara PJ, Pingel K, Yoshitake S;

XX WPI; 1993-133739/16.

XX P-PSDB; AAR34427.

XX Human tissue plasminogen activator single chain form fibrinolytic agent -
 PT comprises thrombin cleavable zymogen stimulating amido lytic activity,
 PT for lysing clots in heart attack and stroke victims and suppressing
 PT fibrin matrix.

XX Example; Fig 8A, 8B, 8C; 22pp; English.

XX A hybrid DNA sequence was constructed which encoded a protein consisting
 CC of the entire amino-terminal portion of t-PA (up to the cysteine at posn.
 CC 261) joined to the serine protease domain of plasminogen beginning at
 CC amino acid 541 (just to the amino-terminal side of the normal activation
 CC site). This hybrid protein was designated "PAP"

XX SQ Sequence 1724 BP; 410 A; 435 C; 508 G; 371 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.24e-111 Length: 1724
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-992-095B-54 (1-207) x AAQ40318 (1-1724)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1033 ATGCACCTTCTGTGGAGGACCTTGATATCCCGAGGTGGGTGTGACTGTGCCCACTGC 1092
 Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
 Db 1093 TTGGAGAGTCCCCCAAGCCCTTTCATCTACAGGTCTATCCGGGTGCACCAAGAAGTG 1152
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1153 AATCTCGNACCGCATGTTCCAGGAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 1212
 Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 1213 AAAGATATGCTTGTCTAAAGCTAAGCAGTCTGCTGCGCTCATCTACGACAAAGTAATCCCA 1272
 Qy 81 AlaCysLeuProSerProAsnTyrlsValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 1273 GCTTGTCTGCATCCCAATATGTTGGTCTGCTGACCGGACCGAATGTTTTCATCACTGGC 1332
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 1333 TGGGGAGAAACCCNAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGTCCCTGTG 1392
 Qy 121 IleGluAsnLysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140

```

Db 1393 ATTGAGAAATTAAGTGTGCAATCGTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 1452
Qy 141 LeuCyAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1453 CTCTGTGCTGGGCAATTTGGCGGAGGACACTGACAGATTGGCAGGGTGACAGTGGAGGTCCT 1512
Qy 161 LeuValCysPheGluLysAspLysValTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1513 CTGGTTTGGCTTCAGAGAGGACAAATACATTTTCAAGGAGTCACTTCTGGGGTCTTGGC 1572
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 1573 TGTGCACGCCCAATAAGCTGGTGTCTATGTTCTGTTTCAAGTTTGTACTTGGATT 1632
Qy 201 GluGlyValMetArgAsnAsn 207
Db 1633 GAGGGAGTGTATGAGAAATAAT 1653

RESULT 4
ACC51086
ID ACC51086 standard; cDNA; 1907 BP.
XX
AC ACC51086;
XX
DT 13-JUN-2003 (first entry)
XX
DE Human Plasminute coding sequence.
XX
KW Human; GENSET; therapeutic; therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200294864-A2.
XX
PD 28-NOV-2002.
XX
PF 06-AUG-2001; 2001WO-IB001715.
XX
PR 25-MAY-2001; 2001US-0293574P.
PR 15-JUN-2001; 2001US-0298698P.
PR 29-JUN-2001; 2001US-0302277P.
PR 13-JUL-2001; 2001US-0305456P.
XX
PA (GEST ) GENSET.
XX
PI Bejanin S, Tanaka H;
XX
DR WPI; 2003-129412/12.
DR P-PSDB; ABR48479.
XX
PT New GENSET polynucleotides and polypeptides, useful for preparing a
PT composition for treating GENSET-related disorders and as reagents in
PT assays to quantitatively determined levels of GENSET expression in
PT biological samples.
XX
PS Claim 1; Page 446-447; 505pp; English.
XX
CC The present invention relates to novel human GENSET coding sequences
CC (ACC51060-ACC51115) and proteins (ABR48453-ABR48508). The GENSET
CC sequences are useful for preparing a composition for treating GENSET-
CC related disorders. They can also be used as markers for tissues in which
CC the corresponding protein is preferentially expressed, as molecular
CC weight markers on Southern gels, as chromosome markers or tags to
CC identify chromosomes, and as reagents in assays to quantitatively
CC determined levels of GENSET expression in biological samples
XX
SQ Sequence 1907 BP; 494 A; 437 C; 427 G; 549 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.42e-111 Length: 1907
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-992-095B-54 (1-207) x ACC51086 (1-1907)
Qy 1 MetHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1044 ATGCACCTTCTGTGAGGCACCTTTGATATCCCGAGAGTGGGTGTGACTGTGCCCCACTGC 1103
Qy 21 LeuGluIysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db 1104 TTGGAGAAGTCCCAAGGCCTTCATCTCACAAGGTCTATCTGGGTGCACACCAAGAAGTG 1163
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1164 AATCTCGAACCGCATGTTCAAGAAATAGAAAGTGTCTAGGCTGTCTTGGAGCCCAACGA 1223
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 1224 AAGATATTGGCTTGTCTAAGCTAAGCAGTCTCTGCCGTCTCATCTGACAAAGTAATCCCA 1283
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 1284 GCTTGTCTGCCATCCCAAAATATGTGTGCTGACCGGACCGAATGTCTTCATCACTGGC 1343
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 1344 TGGGGAGAACCACCAAGGTACTTTTGGAGCTGGCTTCTCAAGAGCCGAGCTCCCTGTG 1403
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1404 ATTGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 1463
Qy 141 LeuCyAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1464 CTCTGTGCTGGGCAATTTGGCCGAGGACACTGACAGTTGCCAGGGTGACAGTGGAGGTCCT 1523
Qy 161 LeuValCysPheGluLysAspLysValTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1524 CTGGTTTGGCTTCAGAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 1583
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 1584 TGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGTTTGTACTTGGATT 1643
Qy 201 GluGlyValMetArgAsnAsn 207
Db 1644 GAGGGAGTGTATGAGAAATAAT 1664

RESULT 5
AAQ40258
ID AAQ40258 standard; DNA; 2296 BP.
XX
AC AAQ40258;
XX
DT 02-AUG-1993 (first entry)
XX
DE Plasmid pN29pt-LPg DNA fragment containing a NotI cleavage site, a
DE selective marker gene, a poxvirus promoter and human plasminogen.
XX
KW Plasmid; cloning; restriction site; ss.
XX
OS Synthetic.
XX
PN AU9221269-A.
XX
PD 04-MAR-1993.
XX
PF 25-AUG-1992; 92AU-00021269.
XX
PR 26-AUG-1991; 91US-00750080.
PR 20-JUL-1992; 92US-00914738.

```

XX PA (IMMO) IMMUNO AG.
 XX PI Dörner F, Scheifflinger F, Falkner FG, Pfeleiderer M;
 XX DR WPT; 1993-126461/16.
 XX PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves direct
 XX PT molecular cloning of modified DNA molecule contg. cytoplasmic DNA virus
 XX PT genome.
 XX PS Claim 67; Page 143-144; 206pp; English.
 XX CC Plasmid pN2 contains a DNA segment having a sequence-specific endonuclease
 CC cleavage site that is unique in the plasmid with a NotI site at each end.
 CC In this plasmid the DNA segment may further comprise a selective marker
 CC gene under transcriptional control of a poxvirus promoter, for instance
 CC pN2-gpta and pN2-gptb. The DNA segment may further comprise a second
 CC poxvirus promoter operatively linked to a DNA sequence comprising a
 CC restriction endonuclease cleavage site, e.g. pN2gpt-S4. pN2gpt-S4 are
 CC comprised of gene expression cassettes suitable for association of an open
 CC reading frame having a translation start codon, with a synthetic promoter
 CC S4 prior to direct molecular transfer into a unique site in vaccinia
 CC virus vector vdrk. Gene expression cassette plasmid pN2gpt-CP9 is for the
 CC expression of human ly8-plasminogen in vaccinia virus vector vdrk
 XX SQ Sequence 2296 BP; 616 A; 597 C; 589 G; 494 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.82e-111 Length: 2296
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-992-095B-54 (1-207) x AAQ04258 (1-2296)

Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1599 ATGCACCTCTGTGAGGACCTTGTATATCCACAGTGGGTGTGACTGTGCTCCACTGC 1658
 Qy 21 LeuGluIysSerProAsgProSerSerTyrrysValIleLeuGlyAlaHisGlnGluVal 40
 Db 1659 TTGGAGAAGTCCCAAGGCCCTTCTCTACAAGGTCTCTCGGTGCACCAAGAAGTG 1718
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1719 AATCTGAACCGCATGTTTCAGGAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 1778
 Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 1779 AAGATATTGCTTGTAAAGCTAAGCAGTCTCTCGCTCATCTGACCAAGTATATCCA 1838
 Qy 81 AlaCysLeuProSerProAsnTyrrValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 1839 GCTTGTCTGCCATCCCAATATATGTGTCTGCTCACCAGCGAATGTTTCATCACTGGC 1898
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 1899 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTCTCAAGGAAGCCCAAGCTCCCTGTG 1958
 Qy 121 IleGluAsnLysValCysAsnArgTyrrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 1959 ATTGAGATAAAGTGTGCATCGCTATAGTTCCTGATGGAAGAGTCCATCCACCGAA 2018
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2019 CTCTGTCTGGGCAATTTGGCCGGAGGCACACAGTTGCCAGGTGCACAGTGGAGGTCT 2078
 Qy 161 LeuValCysPheGluLysAspLysTyrrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2079 CTGGTTTGTCTCGAAGAGGACAAATACATTTTACAAGAGTCACTTCTTGGGGTCTTGGC 2138

Qy 181 CysAlaArgProAsnLysProGlyValTyrrValArgValSerArgPheValThrTrpIle 200
 Db 2139 TGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTCTTCAAGGTTTGTACTTGGATT 2198
 Qy 201 GluGlyValMetArgAsn 207
 Db 2199 GAGGAGTGTGATGAGAAATAAT 2219
 RESULT 6
 AAA89829
 ID AAA89829 standard; DNA; 2296 BP.
 XX AC AAA89829;
 XX DT 26-JAN-2001 (first entry)
 XX DE Plasminogen DNA and S4 promoter.
 XX KW Vaccinia; fowlpox; virus; immune response; HIV; gp-160; gag; gag-pol; ds.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX FN US6103244-A.
 XX PD 15-AUG-2000.
 XX PF 22-MAY-1996; 96US-00651472.
 XX PR 26-AUG-1991; 91US-00750080.
 XX PR 20-JUL-1992; 92US-00914738.
 XX PR 19-DEC-1994; 94US-00358928.
 XX PA (IMMO) IMMUNO AG.
 XX PI Pfeleiderer M, Falkner FG, Scheifflinger F, Dörner F;
 XX DR WPT; 2000-557665/51.
 XX PT Use of modified vaccinia virus and fowlpox virus for generating or
 PT priming an immune response against HIV gp160, HIV Gag and HIV Gag-Pol
 XX proteins.
 XX PS Example 5; Col 107-108; 171pp; English.
 XX CC The present invention relates to the use of modified vaccinia virus and
 CC fowlpox virus for generating or priming an immune response against a
 CC heterologous protein in a vertebrate. Suitable proteins include HIV
 CC proteins such as HIV gp160, HIV Gag and HIV Gag-Pol proteins. The present
 CC sequence was associated with the generation or use of the modified
 CC viruses
 XX SQ Sequence 2296 BP; 622 A; 597 C; 583 G; 494 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.82e-111 Length: 2296
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x AAA89829 (1-2296)

Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1599 ATGCACCTCTGTGAGGACCTTGTATATCCACAGTGGGTGTGACTGTGCTCCACTGC 1658
 Qy 21 LeuGluIysSerProAsgProSerSerTyrrysValIleLeuGlyAlaHisGlnGluVal 40
 Db 1659 TTGGAGAAGTCCCAAGGCCCTTCTCTACAAGGTCTCTCGGTGCACCAAGAAGTG 1718

Qy 41 AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60
 Db |||||
 1719 AATCTCGAACCGCATGTTTCAGAAATAGAGTGCTAGGCTGTTCTTGGAGCCACACGA 1778
 Qy 61 LysAspLeuAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db |||||
 1779 AAGATATTCCTTGCTTAAAGCTAAGCAGTCTCGCGTCATCACTGACAAAGTAATCCCA 1838
 Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 Db |||||
 1839 GCTTGTCTGCCATCCCAATATATGTGTCTGCTGACCGACCGAATGTTTCATCACTGGC 1898
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db |||||
 1899 TGGGAGAAACCCCAAGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCTGTG 1958
 Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db |||||
 1959 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 2018
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db |||||
 2019 CTCGTGTCTGGCAATTTGGCCGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCCT 2078
 Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db |||||
 2079 CTGGTTTCTCGAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 2138
 Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 Db |||||
 2139 TGTGACGCCCAATAAGCTGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2198
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db |||||
 2199 GAGGAGTGTGAGAAATAAT 2219

RESULT 7

AAS12747
 ID AAS12747 standard; DNA; 2296 BP.

XX AAS12747;

XX 21-NOV-2001 (first entry)

XX Plasmid pN2gpt-S4 DNA sequence containing plasminogen cDNA.

XX Cytoplasmic DNA virus; direct molecular cloning; vaccinia virus, insect;
 KW unique restriction endonuclease cleavage site; infectious virion; ds;
 KW helper virus; poxvirus; iridovirus; vertebrate; multiple cloning site;
 KW plasminogen.

XX Synthetic.

XX US6265183-B1.

XX 24-JUL-2001.

XX 19-DEC-1994; 94US-00358928.

XX 26-AUG-1991; 91US-00750080.

XX 20-JUL-1992; 92US-00914738.

XX (BAXT) BAXTER AG.

XX Dorner F, Scheiflinger F, Falkner FG, Pfeleiderer M;

XX WPI; 2001-535006/59.

XX Producing recombinant protein using modified vaccinia viral expression
 PT system, comprises directly cloning heterologous insert encoding protein
 PT into the viral genome into unique restriction endonuclease cleavage site.

XX Example 5; Col 107-110; 172pp; English.

RESULT 8
 ABX15877

XX The invention relates to a method for producing a modified eukaryotic
 CC cytoplasmic DNA virus by direct molecular cloning of a modified DNA
 CC molecule comprising a modified cytoplasmic DNA virus genome such as a
 CC vaccinia virus, containing a heterologous insert encoding a protein. The
 CC method involves molecularly cloning the DNA directly into a host cell via
 CC a unique restriction endonuclease cleavage site, to be packaged into
 CC infectious virions and then recovering them. The host cell is infected
 CC with a helper virus for this purpose. The method is useful for producing
 CC recombinant proteins and modified eukaryotic cytoplasmic DNA viruses such
 CC as poxviruses and iridoviruses found in vertebrates and insects. This
 CC sequence represents the plasmid pN2gpt-S4 containing plasminogen cDNA. It
 CC is used in studies on expression of polypeptides in a vaccinia virus
 CC vector

SQ Sequence 2296 BP; 622 A; 597 C; 583 G; 493 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1.82e-111 Length: 2296
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x AAS12747 (1-2296)

Qy 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1599 ATGCACCTCTGTGAGGACCTTGTATATCCCAAGATGGGTGTGACTGTGCCACTGC 1658
 Qy 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
 Db 1659 TTGGAGAGTCCCCAAGGCCCTTCATCTACAAGGTCTCTGGGTGCACACCAAGAAGTG 1718
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1719 AATCTCGAACCGCATGTTTCAGGAATAGAAAGTGTCTAGGCTGTTCTTGGAGCCACACGA 1778
 Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 1779 AAAGATATTGCCCTGTCTTAAAGCTAAGCAGTCTCGCCGTCATCACTGCAAGTAATCCCA 1838
 Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 1839 GCTTGTCTGCCATCCCAATATATGTTGCTGCTGACCGGACCGAATGTTTCATCACTGGC 1898
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 1899 TGGGAGAAACCCCAAGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCTGTG 1958
 Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 1959 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 2018
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2019 CTCTGTCTGGCATTTGGCCGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCCT 2078
 Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2079 CTGGTTTCTCGAGAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 2138
 Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 Db 2139 TGTGACGCCCAATAAGCTGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2198
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db 2199 GAGGAGTGTGAGAAATAAT 2219

ID ABX15877 standard; DNA; 2430 BP.
 AC ABX15877;
 XX
 XX 15-APR-2003 (first entry)
 XX
 XX DNA encoding human plasminogen protein.
 DE
 DE Anti-angiogenic peptide; human; angp; plasminogen; gene; ds;
 KW gene therapy; angiogenesis; proliferation inhibitor; tumorigenesis;
 KW metastasis; restenosis; retinopathy; atherosclerosis; tumour;
 KW Kaposi's sarcoma; neurofibroma; trachoma; angiogenesis; osteoarthritis;
 KW ocular neovascular disease; age-related macular degeneration;
 KW diabetic retinopathy; rheumatoid arthritis; chronic inflammation;
 KW ulcerative colitis; Crohn's disease; Bartonellosis; haemangioma;
 KW Osler-Weber-Rendu disease; haemorrhagic telangiectasia; contraception.
 KW
 XX Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2430
 FT /*tag= a
 FT /product= "Human plasminogen protein"
 FT /partial
 FT /note= "No stop codon shown"
 FT
 FT sig_peptide 1..5
 FT /*tag= b
 FT mat_peptide 6..2430
 FT /*tag= c
 FT /note= "Mature plasminogen protein"
 FT
 XX US6475784-B1.
 XX
 XX 05-NOV-2002.
 XX
 XX 13-NOV-1998; 98US-00192012.
 XX
 XX 14-NOV-1997; 97US-0066020P.
 XX
 XX (VALE-) VALENTIS INC.
 XX (PFIZ) PFIZER INC.
 XX
 XX Papkoff J;
 XX
 XX WPI; 2003-208837/20.
 DR P-PSDB; ABG75602.
 DR
 XX
 XX New isolated nucleic acid encoding an anti-angiogenic polypeptide that
 PT has kringles 1 - 3 of plasminogen, useful for inhibiting angiogenesis and
 PT endothelial cell proliferation, and for treating e.g. cancer, and
 PT retinopathy.
 XX
 XX Disclosure; Fig 2; 46pp; English.
 XX
 CC This invention relates to an isolated nucleic acid comprising an
 CC expression cassette that comprises a sequence encoding a signal
 CC polypeptide (e.g., an immunoglobulin (Ig/K signal polypeptide) operably
 CC linked to a polynucleotide sequence that encodes an anti-angiogenic
 CC polypeptide, where the protein contains has kringles 1 - 3 of mouse
 CC plasminogen. This sequence may be used in gene therapy, to treat
 CC angiogenesis or as an endothelial cell proliferation inhibitor. It is
 CC also useful for treating angiogenesis-associated conditions such as
 CC tumorigenesis and metastasis and other diseases characterised by
 CC abnormal growth of endothelial cells e.g., restenosis, retinopathies,
 CC atherosclerosis. Tumorigenesis conditions which are treated by the DNA
 CC encoding the Angp protein include malignant solid tumours including
 CC retinoblastomas, Kaposi's sarcoma, etc; and non-malignant tumours
 CC including neurofibromas, trachomas, etc. Other undesirable angiogenesis -
 CC associated conditions include ocular neovascular disease, age-related
 CC macular degeneration, diabetic retinopathy, rheumatoid arthritis,
 CC osteoarthritis, chronic inflammation (including ulcerative colitis,
 CC Crohn's disease, Bartonellosis), atherosclerosis and haemangioma. The
 CC expression cassette is also useful for treating the adverse effects of

CC certain hereditary diseases such as Osler-Weber-Rendu disease and
 CC hereditary haemorrhagic telangiectasia and can be used in contraceptive
 CC methods to block angiogenesis that is required for ovulation and also for
 CC implantation of a blastula after fertilisation. The DNA sequence of the
 CC invention encodes an anti-angiogenic polypeptide that is specific in
 CC inhibiting proliferation of endothelial cells and not of other cell
 CC types, and thus undesired angiogenesis and endothelial cell proliferation
 CC is inhibited, without adversely affecting other cells. The present
 CC sequence represents the DNA sequence encoding the human plasminogen
 CC protein used to generate the anti-angiogenic protein Angp of the
 CC invention
 XX
 SQ Sequence 2430 BP; 676 A; 607 C; 616 G; 531 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.96e-111 Length: 2430
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0
 US-09-992-095B-54 (1-207) x ABX15877 (1-2430)
 Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1810 ATGCACCTCTGTGGAGGCACCTTGATATCCCCAGAGTGGGTGTGACTGTGCCACATGC 1869
 Qy 21 LeuGluLysSerProArgProSerSerTyrIleValIleLeuGlyAlaHisGluVal 40
 Db 1870 TTGGAGAAAGTCCCCAGGCCCTTCATCTCAAGGTCTCTCGGTGCACCAAGAAGTG 1929
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1930 AATCTCGAACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 1989
 Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 1990 AARGATATTGCCTTGTCTAAAGCTAAGCAGTCCTGCCGTCTACCTGACAAAGTAATCCCA 2049
 Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 2050 GCTTGTCTGCCATCCCCAAATTTATGTGTCTGACCGGCCGAGTCTTCTCATCTCCTGGC 2109
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyValAglyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 2110 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCTCTCAAGGAAGCCAGCTCCCTGTG 2169
 Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 2170 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTCCTGAATGGAAGAGTCCAATCCACCGAA 2229
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2230 CTCTGTCTGGCATTTGGCCGGAGGCACCTGACAGTTGCCAGGGTGACAGTGGAGGTCT 2289
 Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2290 CTGGTTGTCTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 2349
 Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 Db 2350 TGTGCAGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2409
 Qy 201 GluGlyValMetArgAsn 207
 Db 2410 GAGGAGGTGATGAGAAATAAT 2430
 RESULT 9
 AAX77711
 ID AAX77711 standard; DNA; 2433 BP.
 XX
 AC AAX77711;

XX 10-AUG-1999 (first entry)
 XX Human plasminogen DNA coding region.
 XX Plasminogen; human; angiotensin; endostatin; gene therapy; vector;
 KW anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
 KW tumour growth; solid tumour; diabetic retinopathy; retina; ss.
 XX Homo sapiens.
 XX WO9926480-A1.
 XX 03-JUN-1999.
 XX 20-NOV-1998; 98WO-US024950.
 XX 20-NOV-1997; 97US-00975424.
 XX (GENE-) GENETIX PHARM INC.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX Leboulich P, Pawliuk RJ, Bachelot T;
 XX WPI; 1999-357696/30.
 XX P-PSDB; AAY08685.
 XX Anti-angiogenic gene therapy vectors.
 XX Disclosure; Fig 5; 83pp; English.
 XX This invention describes a novel viral gene therapy vector comprising a
 CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen from
 CC human or murine angiotensin, human or murine endostatin and angiotensin-
 CC inhibiting fusions and fragments, where the viral vector is sufficiently
 CC attenuated for use in human gene therapy. The products of the invention
 CC have anti-angiogenic, cytostatic, anti-diabetic and ophthalmological
 CC activity. The vector is used in gene therapy for inhibiting tumour growth
 CC in humans harbouring a solid tumour. The vector expresses an anti-
 CC angiogenic polypeptide. An additional use comprises treatment of diabetic
 CC retinopathy, where the anti-angiogenic polypeptide inhibits angiogenesis
 CC in the vicinity of the retina. The vector is administered to cells ex
 CC vivo and then administered to the patient
 XX SQ Sequence 2433 BP; 678 A; 607 C; 616 G; 532 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.96e-111 Length: 2433
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-992-095B-54 (1-207) x AAX77711 (1-2433)
 Qy 1 MethisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1810 ATGCACCTCTGTGGAGGACACCTTGATATCCACAGAGTGGGTGTGACTGCTGCCCATCTGC 1869
 Qy 21 LeuGluIysSerProArgProSerSerTyrIysValIleLeuGlyAlaHisGlnGluVal 40
 Db 1870 TTGGAGAGTCCCCAAGGCCCTTCATCTACAGGTCTCTGGGTGCACACCAAGAGGTG 1929
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1930 AATCTCGAAGCGCATGTTACAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACAGA 1989
 Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspIysValIlePro 80
 Db 1990 AAGATATTGCTTGTCTAAGCTAAGCAGTCTCTGCCGTCTACCTGACCAAGTAAATCCCA 2049
 Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100

Db 2050 GCTTGTCTGCCATCCCAAAATTATGTGGTGCCTGACCGGACCGAAATGTTTCATCCTGGC 2109
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuIysGluAlaGlnLeuProVal 120
 Db 2110 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGGCTTCTCAAGGAAGCCAGCTCCCTGTG 2169
 Qy 121 IleGluAsnIysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 2170 ATTGAGAAATAAAGTGTGCAATCGCTATGAGTTTCTGATGGAAGAGTCCCAATCCACCGAA 2229
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2230 CTCTGTGTGGCATTTGGCCGAGGCACTGACAGTTGCCAGGCTGACAGTGGAGGTCTCT 2289
 Qy 161 LeuValCysPheGluIysAspIysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2290 CTGGTTTGTCTCGAGAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 2349
 Qy 181 CysAlaArgProAsnIysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 Db 2350 TGTGACGCCCAATAGCCTGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2409
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db 2410 GAGGAGTGATGAGAAATAAT 2430
 RESULT 10
 ABN89459
 ID ABN89459 standard; cDNA; 2433 BP.
 XX AC ABN89459;
 DT 02-SEP-2002 (first entry)
 XX Human plasminogen encoding cDNA SEQ ID NO:1.
 XX Human; plasminogen; microplasminogen; miniplasminogen; yeast; vector;
 KW yeast expression vector; cardiant; thrombolytic; cerebroprotective;
 KW fibrin proteolysis; kringle domain; thromboembolic disease;
 KW focal cerebral ischaemic infarction; ischaemic stroke;
 KW arterial thrombotic disease; peripheral arterial occlusive disease;
 KW acute myocardial infarction; gene; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 CDS 1..2433
 FT /*tag= a
 FT /product= "plasminogen"
 XX WO200250290-A1.
 XX 27-JUN-2002.
 XX 20-DEC-2001; 2001WO-BE000217.
 XX 21-DEC-2000; 2000GB-00031196.
 PR 09-JUL-2001; 2001GB-00016690.
 PR 09-JUL-2001; 2001GB-00016702.
 XX (THRO-) THROMB-X NV.
 XX Collen DJ, Nagai N, Laroche Y;
 XX WPI; 2002-500632/53.
 DR P-PSDB; ABB81496.
 XX Novel expression vector for expressing mammalian plasminogen derivatives
 PT in yeast, has nucleotide sequence coding for catalytic domain of
 PT plasminogen and/or coding for kringle domains of plasminogen linked to
 PT promoter.
 XX

PS Claim 6; Page 44-47; 61pp; English.

XX The present invention describes a yeast expression vector (I) comprising
 CC a mammalian nucleotide sequence operably linked to a promoter, where the
 CC mammalian nucleotide sequence codes for the catalytic domain of
 CC plasminogen and further optionally codes for one or more kringle domains
 CC of plasminogen, its mutants or hybrids. (I) has cardant, thrombolytic
 CC and cerebroprotective activities, and can mediate fibrin proteolysis.
 CC (I) can be used useful for treating a thromboembolic disease in a mammal.
 CC The mammalian protein expressed by (I) is useful for treating focal
 CC cerebral ischaemic infarction (ischaemic stroke) or arterial thrombotic
 CC diseases such as peripheral arterial occlusive disease or acute
 CC myocardial infarction. The present sequence encodes human plasminogen,
 CC which is used in the exemplification of the present invention
 XX

SQ Sequence 2433 BP; 678 A; 607 C; 616 G; 532 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.96e-111 Length: 2433
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x ABN89459 (1-2433)

QY 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 DB 1810 ATGCACCTCTGTGGAGGACCTTGTATCCCAAGTGGGTGTACTGTGCCCATGTC 1869
 QY 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGluAlaHisGlnGluVal 40
 DB 1870 TTGGAGAAGTCCCAAGCCCTTCATCTACAAAGTCTCCTGGTGCACACCAAGAAGTG 1929
 QY 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 DB 1930 AATCTCGAACCGCATGTTCCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 1989
 QY 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80
 DB 1990 AAGATATTCCTTGTCTAAGCTAAGCAGTCTCTCCGTCATCTACGAAAGTAATCCCA 2049
 QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 DB 2050 GCTGTCTGCCATCCCAATATATGTGTCTGACCGGACCGAATGTTTCATCACTGGC 2109
 QY 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 DB 2110 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCTGTG 2169
 QY 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 DB 2170 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTCTGAATGGAGAGTCCATCCACCGAA 2229
 QY 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 DB 2230 CTCTGTCTGGCATTTGGCCGAGGACCTGACAGTTGCCAGGTGTGACAGTGGAGGTCT 2289
 QY 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 DB 2290 CTGGTTTGTCTCGAAGAGGACAAATACATTTTACAGAGGTCACTTCTTGGGGTCTTGGC 2349
 QY 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 DB 2350 TGTGACGCCCAATAAGCTGGTGTCTATGTTCTGTGTTCAGGTTTGTACTTGGATT 2409
 QY 201 GluGlyValMetArgAsnAsn 207
 DB 2410 GAGGGAGTGATGAGAAATAT 2430

RESULT 11
 AAQ40319

ID AAQ40319 standard; cDNA; 2679 BP.

XX AAQ40319;
 AC
 XX
 DT 17-AUG-1993 (first entry)
 XX
 DE Sequence of a plasminogen cDNA.
 XX
 DE Zymogen; fibrinolytic activity; cleavage; ss.
 XX
 KW Synthetic.
 OS

XX Key Location/Qualifiers
 FH CDS 22..2454
 FT /*tag= a

XX US5200340-A.

XX 06-APR-1993.

XX 22-MAY-1987; 87US-00053412.

XX 22-MAY-1987; 87US-00053412.

XX (ZYMO) ZYMOGENETICS INC.

XX Foster DC, Mulvihill ER, Ohara PJ, Pingel K, Yoshitake S;

XX MPI; 1993-133739/16.

XX P-PSDB; AAR34428.

XX Human tissue plasminogen activator single chain form fibrinolytic agent -
 PT comprises thrombin cleavable zymogen stimulating amido lytic activity,
 PT for lysing clots in heart attack and stroke victims and suppressing
 PT fibrin matrix.

XX Example; Fig 10A, 10B, 10C; 22pp; English.

XX A lambda phage clone comprising a cDNA sequence encoding plasminogen was
 CC obtained from Dr. Mark Marteen at the University of Washington. The cDNA
 CC was isolated from a human liver library by probing with the partial
 CC sequence of Malinowski et al. The sequence of the complete cDNA and the
 CC encoded amino acid sequence are shown in AAQ40319 and AAR34428

XX Sequence 2679 BP; 750 A; 659 C; 675 G; 595 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.23e-111 Length: 2679
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-992-095B-54 (1-207) x AAQ40319 (1-2679)

QY 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 DB 1831 ATGCACCTCTGTGGAGGACCTTGTATCCCAAGTGGGTGTACTGTGCCCATGTC 1890
 QY 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGluAlaHisGlnGluVal 40
 DB 1891 TTGGAGAAGTCCCAAGCCCTTCATCTACAAAGTCACTCTGGTGCACACCAAGAAGTG 1950
 QY 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 DB 1951 AATCTCGAACCGCATGTTCCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 2010
 QY 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80
 DB 2011 AAGATATTCCTTGTCTAAGCTAAGCAGTCTCTCCGTCATCTACGAAAGTAATCCCA 2070
 QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100

Db 2071 GCTTGTGCTCCATCCCAAAATTAATGCTGCTGACCGGACCGAATGTTTCATCACTGGC 2130
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 2131 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAAGCCAGCTCCCTGTG 2190
 Qy 121 IleGluAenLysValCysAenArgTyrGluPheLeuAenGlyArgValGlnSerThrGlu 140
 Db 2191 ATTGAGAAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAGAGTCCATCCACCGAA 2250
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2251 CTCTGTGTGGGCAATTTGGCCGAGGACCTGACAGTTGCCAGGCTGACAGTGGAGGTCTCT 2310
 Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2311 CTGGTTTGTCTCGAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 2370
 Qy 181 CysAlaArgProAenLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 Db 2371 TGTGACGCCCCCAATAGCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2430
 Qy 201 GluGlyValMetArgAenAen 207
 Db 2431 GAGGAGTGATGAGAAATAAT 2451

RESULT 12

AAX35376
 ID AAX35376 standard; DNA; 2732 BP.

XX AC AAX35376;

XX DT 16-JUL-1999 (first entry)

XX DE SEQ ID 51 of W09916889.

XX KW Angiostatin; endostatin; interferon; thrombospondin;
 XX KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
 XX KW anti-tumor; multifunctional protein; angiogenic-mediated disease; cancer;
 XX KW diabetic retinopathy; macular degeneration; arthritis;
 XX KW tumor cell production; ss.

XX OS Homo sapiens.

XX XX W09916889-A1.

XX XX 08-APR-1999.

XX XX 30-SEP-1998; 98WO-US020464.

XX XX 01-OCT-1997; 97US-0060609P.

XX XX (SEAR) SEARLE & CO G D.

XX PI Bolanowski MA, Caparon MH, Casperson GP, Gregory SA, Klein BK;
 XX PI McKearn JP;

XX DR WPI; 1999-255098/21.

XX FT New multifunctional proteins useful for treating angiogenic-mediated
 XX FT diseases.

XX PS Disclosure; Page 85-86; 121pp; English.

XX CC The specification describes multifunctional proteins which comprise
 CC combinations of angiotatin, endostatin, interferon, thrombospondin,
 CC interferon-inducible protein and platelet factor 4, and have anti-
 CC angiogenic and/or anti-tumor activity. The multifunctional protein may
 CC exhibit useful properties such as having similar or greater biological
 CC activity when compared to a single factor or by having improved half-life
 CC or decreased adverse side effects, or a combination of these properties.
 CC The proteins can be used for treating an angiogenic-mediated disease,

CC e.g. cancer, diabetic retinopathy, macular degeneration, or arthritis.
 CC They can also be used for inhibiting the production of tumor cells
 CC (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric,
 CC colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma)
 CC in a patient and for inhibiting tumor growth. The present sequence is
 CC used in the course of the invention

XX SQ Sequence 2732 BP; 757 A; 667 C; 690 G; 618 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,29e-111 Length: 2732
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-992-095B-54 (1-207) x AAX35376 (1-2732)

Qy 1 MetHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1864 ATGCACCTTCTGTGGAGGACCTTGATATCCAGAGTGGGTGTGACTGTGCTGCCACTGC 1923
 Qy 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
 Db 1924 TTGGAGAAAGTCCCAAGGCCTTCATCTCAAGGTCACTCTGGGTGCACACCAAGAAAGTG 1983
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1984 AATCTCGAACCGCATGTTCAGGAAATAGAGTGTCTAGGCTGTTCCTTGGAGCCACACGA 2043
 Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 2044 AAGATATTGCTTGTCTTAAGCTAAGCAGTCTCTGCCGTCACTGACAAAGTAATCCCA 2103
 Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 2104 GCTTGTCTGCATCCCAAAATATGTGTGTCGTCGACCGGACCGAAATGTTTCATCACTGGC 2163
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 2164 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCAGTCCCTGTG 2223
 Qy 121 IleGluAenLysValCysAenArgTyrGluPheLeuAenGlyArgValGlnSerThrGlu 140
 Db 2224 ATTGAGAAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAGAGTCCATCCACCGAA 2283
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2284 CTCTGTGTGGGCAATTTGGCCGAGGACCTGACAGTTGCCAGGCTGACAGTGGAGGTCTCT 2343
 Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2344 CTGGTTTGTCTCGAGAGGACAAATACATTTTAAAGGAGTCACTTCTTGGGTCTTGGC 2403
 Qy 181 CysAlaArgProAenLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 Db 2404 TGTGACGCCCCCAATAGCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2463
 Qy 201 GluGlyValMetArgAenAen 207
 Db 2464 GAGGAGTGATGAGAAATAAT 2484

RESULT 13

ABN81696

ID ABN81696 standard; DNA; 2732 BP.

XX AC ABN81696;

XX DT 03-SEP-2002 (first entry)

XX DE Human plasminogen encoding DNA.

XX XX

KW Human; plasminogen; lys plasminogen; glu plasminogen; A61; P22;
 KW vascular endothelial; cell proliferation; anti-angiogenic; cancer;
 KW rheumatoid arthritis; Crohn's disease; diabetic retinopathy; anti-cancer;
 KW cytostatic; gynaecological; neuroprotective; antipsoriatic;
 KW antiarteriosclerotic; dermatological; antirheumatic; antiarthritic;
 KW antiinflammatory; antidiabetic; ophthalmological; immunosuppressive;
 KW cardiant; vulnerary; vasotropic; anti-tumour; gene; ds.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 XX CDS 1..2487
 XX /*tag= a
 XX /product= "plasminogen"
 XX sig_peptide 55..111
 XX /*tag= b
 XX mat_peptide 112..2484
 XX /*tag= c
 XX /product= "glu plasminogen"
 XX mat_peptide 343..2484
 XX /*tag= d
 XX /product= "lys plasminogen"
 XX mat_peptide 343..1524
 XX /*tag= f
 XX /product= "A61 isoform"
 XX mat_peptide 343..1515
 XX /*tag= e
 XX /product= "A61 isoform"
 XX mat_peptide 343..651
 XX /*tag= g
 XX /product= "P22"

XX WO200244328-A2.

XX 06-JUN-2002.

XX 28-NOV-2001; 2001WO-US044515.

XX 28-NOV-2000; 2000US-0253725P.

XX (WAIS/) WAISMAN D M.
 XX (KASS/) KASSAM G.
 XX (KWON/) KWON M.

XX Waisman DM, Kassam G, Kwon M;

XX WPI; 2002-527706/56.
 XX P-PSDB; ABB83795.

XX Novel naturally occurring fragment A16 or P22 of plasminogen, useful for
 XX treating for anti-angiogenic treatment of a mammal suffering from cancer
 XX and inhibiting proliferation of vascular endothelial cells.

XX Claim 26; Fig 2; 89pp; English.

XX The invention relates to an isolated naturally occurring fragment A61 or
 XX p22 polypeptide (I) of plasminogen or an isolated polypeptide comprising
 XX 103, 391 or 394 contiguous amino acids of amino terminal of plasminogen
 XX (ABB83795). (I) has vascular endothelial cell proliferation inhibitor
 XX activity. (I) is useful for anti-angiogenic treatment of a mammal
 XX suffering from cancer, acoustic neuromas, neurofibromas, trachomas,
 XX pyogenic granulomas, telangiectasias, psoriasis, scleroderma,
 XX atherosclerosis, rheumatoid arthritis, Crohn's disease, endometriosis,
 XX adiposity, pyogenic granuloma, rubeosis, and diabetic retinopathy,
 XX retinopathy of prematurity, neovascular glaucoma, retrolental
 XX fibroplasia, graft rejection, myocardial angiogenesis, plaque
 XX neovascularization, haemophilic joints, angiofibroma and wound
 XX granulation and inhibiting proliferation of vascular endothelial cells,
 XX in combination with an anti-cancer agent e.g. methotrexate, mitozantrone,
 XX paclitaxel, vinblastine, 5-fluorouracil, cisplatin, leucovorin,
 XX cyclophosphamide and oncolytic virus

XX Sequence 2732 BP; 758 A; 669 C; 688 G; 617 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,29e-111 Length: 2732
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x ABB81696 (1-2732)

Qy	1	MetHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys	20
Db	1864	ATGCACCTCTGTGTGGAGGCACCTTGATATCCCAAGTGGGTGTGACTGTGCCACACGTC	1923
Qy	21	LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal	40
Db	1924	TTGGAGAAAGTCCCCNAGGCTTCATCTACAGGTCTCTGGGTGCACACCAAGAGTG	1983
Qy	41	AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg	60
Db	1984	AATCTCGAACCGCATGTTTCAGGAAATAGAAAGTGTCTAGGCTGTCTTGGAGCCACACGA	2043
Qy	61	LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro	80
Db	2044	AAAGATATTGCTTGTCTAAAGCTAAGCAGTCTCTGCCGTCTACTGACAAAGTAATCCCA	2103
Qy	81	AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly	100
Db	2104	GCITGTCTGCCATCCCAATTTATGTGTCTGACCGACCGAATGTTTTCATCCTGGC	2163
Qy	101	TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal	120
Db	2164	TGGGAGAAACCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG	2223
Qy	121	IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu	140
Db	2224	ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA	2283
Qy	141	LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro	160
Db	2284	CTCTGTCTGGGCATTTGGCGGAGGCACCTGACAGTTGCCAGGGTGACAGTGGAGGTCT	2343
Qy	161	LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly	180
Db	2344	CTGGTTTGTCTCGAGAAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGTCTTGGC	2403
Qy	181	CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle	200
Db	2404	TGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTTCAAGGTTTGTACTTGGATT	2463
Qy	201	GluGlyValMetArgAsnAsn	207
Db	2464	GAGGAGTGTATGAGAAATAAT	2484

RESULT 14

AAD56132
 ID AAD56132 standard; DNA; 2732 BP.

XX AC AAD56132;

XX DT 07-AUG-2003 (first entry)

XX DE Human plasminogen DNA.

XX Osteoarthritis; rheumatoid arthritis; plasmin; plasminogen; human;
 KW urokinase-type plasminogen activator; uPA; degenerative joint disease;
 KW spondyloarthritis; antisense-therapy; antibody therapy; osteopathic;
 KW urokinase-type plasminogen activator receptor; psoriatic arthritis;
 KW plasminogen-activator inhibitor type 1; PAI-1; uPAR; gene; ds.
 XX OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 CDS 55..2487
 FT /*tag= a
 FT /product= "Human plasminogen protein"
 XX
 XX WO2003033009-A2.
 XX
 XX PD 24-APR-2003.
 XX
 XX PF 10-JUL-2002; 2002WO-IB005797.
 XX
 XX PR 10-JUL-2001; 2001US-0304461P.
 PR 10-JUL-2001; 2001US-0304490P.
 PR 13-JUL-2001; 2001US-0305182P.
 XX
 XX FA (ONVI-) ONVIO AB.
 XX
 XX PT NY T, Holmdahl R, Li J;
 XX WPI; 2003-393477/37.
 DR P-PSDB; RAE37127.
 XX
 XX Treating or preventing arthritis e.g. osteoarthritis, rheumatoid
 PT arthritis in mammals, by administering inhibitors of plasmin,
 PT plasminogen, urokinase-type plasminogen activator or plasminogen-
 PT activator inhibitor type 1.
 XX
 XX Disclosure; Page 61-63; 85pp; English.
 XX
 XX The invention relates to a method for treating or preventing arthritis
 CC e.g. osteoarthritis, rheumatoid arthritis in mammals, by administering
 CC inhibitors of plasmin, plasminogen, urokinase-type plasminogen activator
 CC (uPA), plasminogen-activator inhibitor type 1 (PAI-1) and urokinase-type
 CC plasminogen activator receptor (uPAR). The method is useful for treating
 CC or preventing arthritis caused by degenerative joint disease, preferably
 CC rheumatoid arthritis, psoriatic arthritis, infectious arthritis, juvenile
 CC rheumatoid arthritis, osteoarthritis and spondyloarthropathies in a
 CC mammal, especially a human. It is also useful for identifying agents for
 CC treating or preventing arthritis in a mammal and it is also useful in
 CC antisense-therapy and antibody therapy. The present sequence is human
 CC plasminogen DNA
 XX
 XX SQ Sequence 2732 BP; 758 A; 669 C; 688 G; 617 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. NO.: 2.29e-111 Length: 2732
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0
 US-09-992-095B-54 (1-207) x AAD56132 (1-2732)
 QY 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 DB 1864 ATGCATCTCTGTGGAGCACCCTTGATATCCAGAGTGGGTGTGCTGCTGCCCATGTC 1923
 QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
 DB 1924 TTGGAGAAGTCCCAAGGCTTCATCTACAGGTCTGCTGAGTGCACCAAGAGTGT 1983
 QY 41 AsnLeuGluProHisValGlnGlnIleGluValSerArgGluLeuProThrArg 60
 DB 1984 AATCTCGAACCGCATGTTTCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 2043
 QY 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80
 DB 2044 AAGATATGCTTGTCTTAAGCTAAGCAGTCTCCGCTCATCTGACAAAGTATATCCCA 2103
 QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
 DB 2104 GCTTGTCTGCCATCCCAAAATTATGTGTGCTGCTGACCGGACCGAATGTTTCATCACTGGC 2163

QY 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 DB 2164 TGGGGAGAAACCCAGGTACTTTTGGAGCTGGGCTTCTCAAGNAGCCAGCTCCCTGTG 2223
 QY 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 DB 2224 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 2283
 QY 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 DB 2284 CTCTGTCTGGGATTTGGCCGAGGCACTGACAGTTTGCAGGCTGACAGTGGAGGTCT 2343
 QY 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 DB 2344 CTGTTTCTTCGAGAGGCAATATATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 2403
 QY 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
 DB 2404 TGTGCACGCCCAATTAAGCTGTGCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2463
 QY 201 GluGlyValMetArgAsnAsn 207
 DB 2464 GAGGAGTGTGATGAGAAATAT 2484
 RESULT 15
 ABX74464
 ID ABX74464 standard; cDNA; 2732 BP.
 AC ABX74464;
 AT 21-MAR-2003 (first entry)
 DE Human cDNA sequence #7 down-regulated in CC-RCC patients.
 KW Human; microarray; solid surface; immobilised probe; CC-RCC;
 KW differential expression profile; aggressive CC-RCC tumour type;
 KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
 KW gene expression profiling; tumour tissue; gene; ss.
 OS Homo sapiens.
 XX WO200279411-A2.
 XX 10-OCT-2002.
 XX 29-MAR-2002; 2002WO-US009576.
 XX 29-MAR-2001; 2001US-0279411P.
 XX (VAND-) VAN ANDEL INST.
 XX Haab B, Rhodes D, Teh BT, Takashi M;
 WPI; 2003-040679/03.
 PT New microarray, comprising a matrix of cDNA probe from a set of probes
 PT immobilized to a solid surface in predetermined order, useful in the
 PT prognosis of patients with clear cell renal carcinoma.
 PS Example 3; Page 138-139; 179pp; English.
 CC The present invention relates to a microarray comprising a matrix of at
 CC least one cDNA probe from a set of probes immobilised to a solid surface
 CC in a predetermined order, where a row of pixels corresponds to replicates
 CC of one distinct probe from the set. The probes are complementary to
 CC nucleic acid sequences that are expressed differentially in aggressive as
 CC compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
 CC and that hybridise to the probes under high stringency conditions. The
 CC microarray is useful for the prognosis of patients with CC-RCC, wherein
 CC aggressive and non-aggressive CC-RCC tumour types are characterised by
 CC differential expression profiles of genes that hybridise with one or more
 CC probes immobilised on the microarray. The arrays are useful for gene

CC expression profiling of tumour and normal tissues. The present sequence
CC represents a human cDNA sequence down-regulated in CC-RCC patients
XX
SQ Sequence 2732 BP; 760 A; 668 C; 687 G; 617 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,29e-111 Length: 2732
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-992-095B-54 (1-207) x ABX74464 (1-2732)

```
Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
   |||||
Db 1864 ATGCACITCTCTGGAGGACCTTGATATCCCGAGAGTGGGTGTGACTGCTGCCACATGC 1923
   |||||

Qy 21 LeuGluLysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
   |||||
Db 1924 TTGGAGAAAGTCCCAAGGCTTCATCTACCAAGGTCTCTGGGTGCACACCAAGAAGTG 1983
   |||||

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
   |||||
Db 1984 AATCTCGAACCGCATGTTTCAGGAATAGAAAGTGTCTAGGCTGTCTTGAGGCCACACAGA 2043
   |||||

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
   |||||
Db 2044 AAGATATTGCTTGTATAGCTAAGCAGTCTCTGCCGTCTCATCTGACAAAGTAAATCCCA 2103
   |||||

Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaSerArgThrGluCysPheIleThrGly 100
   |||||
Db 2104 GCTTGTCTGCCATCCCAAAATATGTGTGTCTGACCGGACCGAATGTTTCATCACTGGC 2163
   |||||

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
   |||||
Db 2164 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGGCCAGCTCCCTGTG 2223
   |||||

Qy 121 IleGluAsnLysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140
   |||||
Db 2224 ATTGAGATAAAGTGTGCAATCGCTATGAGTTCGTGAATGGAGAGTCCATCCACCGAA 2283
   |||||

Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
   |||||
Db 2284 CTCTGTCTGGGCATTTGGCCGGAGGCACCTCACAGTTGCCAGGTGACAGTGGAGGTCT 2343
   |||||

Qy 161 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180
   |||||
Db 2344 CTGGTTTGTCTTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 2403
   |||||

Qy 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTriple 200
   |||||
Db 2404 TGTGACGCCCAATAGCCCTGGTGTCTATGTTCTGTTTCAGGTTTGTACTTGGATT 2463
   |||||

Qy 201 GluGlyValMetArgAsnAsn 207
   |||||
Db 2464 GAGGAGTGTATGAGAAATAAT 2484
   |||||
```

Search completed: September 13, 2004, 23:33:08
Job time : 407 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 13, 2004, 22:53:55 ; Search time 78 Seconds
(without alignments)
1472.754 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113

Sequence: 1 MHFCGGTLPSPVLTAAHC.....GVTVRSRFTWIEGVNRN 207

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPIO_spool_1/US0992095/runat_10092004_104839_4960/app_query.fasta_1.391
-DB=issued Patents NA -QFW=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCH=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0992095 @CGN 1.1 69 @runat_10092004_104839_4960 -NCPUG=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1113	100.0	1724	6	US0992095
2	1113	100.0	2296	1	Patent No. 5200340-5
3	1113	100.0	2296	3	Sequence 18, Appl
4	1113	100.0	2296	3	Sequence 18, Appl
5	1113	100.0	2430	4	Sequence 4, Appl
6	1113	100.0	2679	6	Sequence 18, Appl
7	1113	100.0	2753	1	Patent No. 5200340
8	1112	99.9	2497	1	Sequence 1, Appl
9	1112	99.9	2497	1	Sequence 12, Appl
10	1112	99.9	2497	2	Sequence 12, Appl
11	1112	99.9	2497	3	Sequence 12, Appl
12	1112	99.9	2497	3	Sequence 12, Appl

13	1106	99.4	690	1	US-08-379-621-1	Sequence 1, Appl
14	1106	99.4	690	2	US-08-889-078-1	Sequence 1, Appl
15	444	39.9	1615	4	US-09-820-002-1	Sequence 1, Appl
16	444	39.9	1783	3	US-09-510-738A-188	Sequence 188, App
17	444	39.9	1783	4	US-09-861-966-188	Sequence 188, App
18	444	39.9	2383	4	US-09-742-703-3	Sequence 3, Appl
19	440.5	39.6	2413	3	US-09-518-046-1	Sequence 1, Appl
20	437	39.3	1130	4	US-09-387-375-8	Sequence 8, Appl
21	437	39.3	1613	4	US-09-387-375-1	Sequence 30, Appl
22	434	39.0	980	4	US-09-023-942A-30	Sequence 1, Appl
23	434	39.0	1110	4	US-09-386-653A-1	Sequence 1, Appl
24	434	39.0	1130	4	US-09-386-653A-8	Sequence 8, Appl
25	433	38.9	1605	2	US-09-000-846-1	Sequence 1, Appl
26	428	38.5	1212	4	US-09-620-312D-431	Sequence 431, App
27	424.5	38.1	959	4	US-09-023-942A-25	Sequence 25, Appl
28	423.5	38.1	2581	5	PCT-US94-00616-1	Sequence 1, Appl
29	423.5	38.1	2581	5	PCT-US94-00616-1	Sequence 1, Appl
30	420	37.7	2544	3	US-09-518-046-3	Sequence 3, Appl
31	419	37.6	933	4	US-09-023-942A-29	Sequence 29, Appl
32	419	37.6	1031	2	US-08-978-404B-1	Sequence 1, Appl
33	409.5	36.8	897	2	US-08-956-267A-1	Sequence 1, Appl
34	405.5	36.4	2416	3	US-09-361-416-1	Sequence 1, Appl
35	403	36.2	1094	4	US-09-023-942A-3	Sequence 3, Appl
36	402.5	36.2	825	3	US-09-120-582-1	Sequence 1, Appl
37	402	36.1	1103	4	US-09-386-642-59	Sequence 59, Appl
38	401	36.0	1081	3	US-09-008-271A-15	Sequence 15, Appl
39	401	36.0	1100	4	US-09-907-794A-256	Sequence 256, App
40	401	36.0	1100	4	US-09-905-125A-256	Sequence 256, App
41	401	36.0	1100	4	US-09-902-775A-256	Sequence 256, App
42	400	35.9	1100	4	US-09-023-942A-5	Sequence 5, Appl
43	400	35.9	1219	2	US-08-978-404B-7	Sequence 7, Appl
44	400	35.9	3147	2	US-09-027-337-1	Sequence 1, Appl
45	400	35.9	3147	4	US-09-644-600-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

5200340-5

; Patent No. 5200340

; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,

; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI

; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN

; ACTIVATORS

; NUMBER OF SEQUENCES: 34

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/53,412

; FILING DATE: 22-MAY-1987

; SEQ ID NO:5:

; LENGTH: 1724

5200340-5

Alignment Scores:

Pred. No.: 3,79e-130 Length: 1724
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x 5200340-5 (1-1724)

Qy	1	MetHisPheCysGlyThrLeuLeuSerProGluTrpValLeuThrAlaAlaHisCys	20
Db	1033	ATGCACCTCTGTGGAGGCACCTTGATATCCAGAGTGGGTGTGACTGTGCTGCCTGC	1092
Qy	21	LeuGluLysSerProArgProSerSerTyrLysValleLeuGlyAlaHisGlnGluVal	40
Db	1093	TTGGAGAGTCCCAAGGCTTCATCTACAGTCTATCTGGTGCACCAAGAGTG	1152
Qy	41	AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg	60
Db	1153	AATCTCGAACCGCATGTTCAGGAATAGAGTGTCTAGGCTGTTCCTGGAGCCACACGA	1212

QY 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80
Db 1213 AAAGATATTCCTTGCCTAAAGCTAAGCAGCTCTGCCCTCATCTGACAAAGTAATCCCA 1272
QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 1273 GCTTGTCTGCCATCCCAATATGTTGGTGGCTGACCGGACCGAATGTTTCATCCTGGC 1332
QY 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 1333 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 1392
QY 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1393 ATTGAGAAATGAAGTGCATCGCTATGAGTTCTGAATGAAGAGTCCATCCACCGAA 1452
QY 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1453 CTCTGTCTGGCATTTGGCGGAGGCACTGACAGTTGCCAGGTGACAGTGGAGGTCTT 1512
QY 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1513 CTGTTTGTCTCGAAGAGCAATATACATTTTCAAGGAGTCACTTCTTGGGCTTGGC 1572
QY 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 1573 TGTGACGCCCCAATAAGCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1632
QY 201 GluGlyValMetArgAsnAsn 207
Db 1633 GAGGAGTGTAGATAAATAAT 1653

RESULT 2

US-07-750-080A-18
; Sequence 18, Application US/07750080A
; Patent No. 5445953
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF A MODIFIED
; TITLE OF INVENTION: EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07750,080A
; APPLICATION NUMBER: 19910826
; FILING DATE: 19910826
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/106 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2296 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: pNZgpt-LPg (Fig. 5.3)
US-07-750-080A-18
Alignment Scores:
Pred. No.: 5 88e-130 Length: 2296
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-992-095B-54 (1-207) x US-07-750-080A-18 (1-2296)
QY 1 MetHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1599 ATGCACCTTCTGTGGAGCACCTTGATATCCAGAGTGGGTGTGACTGTGCTGCCACTGC 1658
QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnVal 40
Db 1659 TTGGAGAGTCCCCAAGGCCCTTCATCTCAAGGTCTCTGGGTGCACACCAAGAGTG 1718
QY 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1719 AATCTCGAACCGCATGTTCCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 1778
QY 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 1779 AAAGATATTGCTTGTCTAAAGCTTAAGCAGTCTCTCCGCTCATCTGACAAAGTAATCCCA 1838
QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 1839 GCTTGTCTGCCATCCCCAAATATGTTGGTGTGCTGACCGGACCGAATGTTTCATCCTGGC 1898
QY 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 1899 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAGCCAGCTCCCTGTG 1958
QY 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1959 ATTGAGATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 2018
QY 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2019 CTCTGTCTGGGCATTTGGCCGAGGCACTGACAGTTGCCAGGTGTACAGTGGAGGTCT 2078
QY 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2079 CTGTTTGTCTTCGAGAAGGACAATATACATTTTACAGGAGTCACTTCTTGGGTCTTTGGC 2138
QY 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 2139 TGTGCAAGCCCAATAAGCTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2198
QY 201 GluGlyValMetArgAsnAsn 207
Db 2199 GAGGAGTGTAGATAAATAAT 2219
RESULT 3
US-08-651-472-18
; Sequence 18, Application US/08651472
; Patent No. 6103244
; GENERAL INFORMATION:
; APPLICANT: DORNER, Friedrich
; APPLICANT: SCHEIFLINGER, Friedrich
; APPLICANT: FALKNER, Falko Gunter
; APPLICANT: PELEIDERER, Michael
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
; TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
; TITLE OF INVENTION: (HIV-1) ANTIGENS
; NUMBER OF SEQUENCES: 95


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,472
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,738
FILING DATE: 20-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,080
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/166/IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
IMMEDIATE SOURCE:
CLONE: pN2gpt-LPg
US-08-651-472-18

Alignment Scores:
Pred. No.: 5.88e-130 Length: 2296
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-651-472-18 (1-2296)

Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1599 ATGCACCTCTGTGGAGGCACCTTGATATCCCGAGGTGGGTGTGACTGTGCCCACTGC 1658
Qy 21 LeuGluIysSerProArgProSerSerTyrylValIleLeuGlyAlaHisGlnGluVal 40
Db 1659 TTGAGAAAGTCCCAAGGCCTTCATCTTACAAAGTTCATCTGGGTGACACCAAGAAGTG 1718
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1719 AATCTCGAACCGCATGTTCCAGAAATAGAGTGTAGGCTGTCTTGGAGCCCAACAGA 1778
Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProIleValIleThrAspLysValIlePro 80
Db 1779 AAAGATATTCCTTGCTAAAGCTAAGCAGTCTCGCGTCATCACTGACAAAGTAATCCCA 1838
Qy 81 AlaCysLeuProSerProAsnTyrylValIleAlaAspArgThrGluCysPheIleThrGly 100
Db 1839 GCITTTGCTGCCATCCCAAAATATATGTGGTCTGCTACCGGACCGAATGTTTCATCAGTGC 1898
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuIysGluAlaGlnLeuProVal 120

```

```

Db 1899 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCCACTCCCTGTG 1958
Qy 121 IleGluAsnLysValCysAsnArgTyrylGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1959 ATTGAGAAATAAGTGTCGAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 2018
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2019 CTCTGTCTGGCAATTTGGCGGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCTCT 2078
Qy 161 LeuValCysPheGluLysAspLysTyrylIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2079 CTGTTTGTCTTCGAGAAGGACAAATACATTTTACAGAGTCACTTCTTGGGGTCTTGGC 2138
Qy 181 CysAlaArgProAsnLysProGlyValTyrylValArgValSerArgPheValThrTrpIle 200
Db 2139 TGTGCAGCCCAATAGCCGTGTCTATGTTCTGTGTTTCAAGGTTTGTGTACTTGGATT 2198
Qy 201 GluGlyValMetArgAsnAsn 207
Db 2199 GAGGAGTGTGAGAAATAAT 2219

RESULT 4
US-08-358-928-18
; Sequence 18, Application US/08358928
; Patent No. 6265183
; GENERAL INFORMATION:
; APPLICANT: DORNER, Friedrich
; APPLICANT: SCHEIFLINGER, Friedrich
; APPLICANT: FALKNER, Falko Gunter
; APPLICANT: PFLEIDERER, Michael
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
; TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,928
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,738
; FILING DATE: 20-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/750,080
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/166/IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid;

```

DESCRIPTION: Synthetic DNA oligonucleotide

IMMEDIATE SOURCE:

CLONE: pN2gpt-LP

US-08-358-928-18

Alignment Scores:

Pred. No.: 5.88e-130 Length: 2296
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-358-928-18 (1-2296)

Qy 1 MethHisPheCysGlyGlyThrLeuLeuSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1599 ATGCACCTTCTGGAGGCACCTTGATATCCACAGTGGGTGTGACTGTGCCCACTGC 1658
Qy 21 LeuGluLysSerProArgProSerSerTyrlsValleLeuGlyAlaHisGlnGluVal 40
Db 1659 TTGGAGAAGTCCCAAGGCCCTTCATCTACAGGTCACTCTGGGTGCACCAAGAAGTG 1718
Qy 41 AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60
Db 1719 AATCTCGAACCGCATGTTACAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACGA 1778
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValleThrAspLysVallePro 80
Db 1779 AAAGATATTTGCTTGTCTAAAGCTAAGCAGTCTCTGCCGTCTACCTGACAAAGTAATCCCA 1838
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 1839 GCTTGTCTGCATATCCCAATATGTTGCTGCTGACCGACCGAATGTTTCACTTGGC 1898
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 1899 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 1958
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1959 ATTGAGAATAAAGTGTGCAATTCGCTATGAGTTTCTGAATGGAAGAGTCCCAATCCACCGAA 2018
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2019 CTCGTGCTGGGCATTTGGCCGAGGACCTGACAGTTGCCAGGGTGACAGTGGAGGTCT 2078
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2079 CTGGTTTCTCGAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 2138
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 2139 TGTGACGCCCAATAAGCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2198
Qy 201 GluGlyValMetArgAsnAsn 207
Db 2199 GAGGAGTGATGAGAAATAAT 2219

RESULT 5

US-09-192-012-4

Sequence 4, Application US/09192012A

Patent No. 6475784

GENERAL INFORMATION:

APPLICANT: Papkoff, Jackie

APPLICANT: Megabios Corporation

APPLICANT: Pfizer, Inc.

TITLE OF INVENTION: Inhibition of Angiogenesis by Delivery of Nucleic Acids

TITLE OF INVENTION: Encoding Anti-Angiogenesis Polypeptides

FILE REFERENCE: 018484-00011005

CURRENT APPLICATION NUMBER: US/09/192,012A

CURRENT FILING DATE: 1998-11-13

EARLIER APPLICATION NUMBER: US 60/066,020

EARLIER FILING DATE: 1997-11-14

NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4

LENGTH: 2430

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2430)

OTHER INFORMATION: human plasminogen

US-09-192-012-4

Alignment Scores:

Pred. No.: 6.42e-130 Length: 2430
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-192-012-4 (1-2430)

Qy 1 MethHisPheCysGlyGlyThrLeuLeuSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1810 ATGCACCTTCTGGAGGCACCTTGATATCCACAGTGGGTGTGACTGTGCCCACTGC 1869
Qy 21 LeuGluLysSerProArgProSerSerTyrlsValleLeuGlyAlaHisGlnGluVal 40
Db 1870 TTGGAGAAGTCCCAAGGCCCTTCATCTACAGGTCTATCTCTGGGTGCACCAAGAAGTG 1929
Qy 41 AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60
Db 1930 AATCTCGAACCGCATGTTACAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACGA 1989
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValleThrAspLysVallePro 80
Db 1990 AAAGATATTTGCTTGTCTAAAGCTAAGCAGTCTCTGCCGTCTACCTGACAAAGTAATCCCA 2049
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 2050 GCTTGTCTGCATATCCCAATATGTTGCTGCTGACCGACCGAATGTTTCACTTGGC 2109
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 2110 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 2169
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 2170 ATTGAGAATAAAGTGTGCAATTCGCTATGAGTTTCTGAATGGAAGAGTCCCAATCCACCGAA 2229
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2230 CTCGTGCTGGGCATTTGGCCGAGGACCTGACAGTTGCCAGGGTGACAGTGGAGGTCT 2289
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2290 CTGGTTTCTCGAGAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 2349
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 2350 TGTGACGCCCAATAAGCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2409
Qy 201 GluGlyValMetArgAsnAsn 207
Db 2410 GAGGAGTGATGAGAAATAAT 2430

RESULT 6

5200340-7

Patent No. 5200340

APPLICANT: FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,

PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI

TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN

```

;ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO.7:
; LENGTH: 2679
5200340-7

Alignment Scores:
Pred. No.: 7.46e-130 Length: 2679
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x 5200340-7 (1-2679)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1831 ATGCACCTTCTGTGGAGGCACCTTGATATCCCGAGAGTGGGTGTTGACTGCTGCCACTGC 1890

Qy 21 LeuGluYsSerProArgProSerSerTyRlyValIleLeuGlyAlaHisGluVal 40
Db 1891 TTGGAGAAGTCCCAAGGCCCTTCTCTACAGGTGCATCTGGGTGGCACCAAGAAGTG 1950

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1951 AATCTCGNACCGCATGTCAGAAATAGATGTCTAGGCTGTTCTTGAGCCACACGA 2010

Qy 61 LysAspIleAlaLeuLeuLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 2011 AAAGATATTGCTTGTAAAGCTAAGCAGTCTGCGGTGCATCACTGACAAAGTAATCCA 2070

Qy 81 AlaCysLeuProSerProAsnTy-ValValAlaAspArgThrGluCysPheIleThrGly 100
Db 2071 GCTTGTCTGCCATCCCAAAATATATGTGTGCTGCTGACCGACCGAATGTTTCATCACTGC 2130

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 2131 TGGGGAGNAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 2190

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 2191 ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGAAGAGTCCAATCCACGAA 2250

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2251 CTCTGTGCTGGGCATTTGGCCGGAGGCACCTGACAGTTGCCAGGGTGACAGTGGAGTCT 2310

Qy 161 LeuValCysPheGluLysAspLysTyrlleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2311 CTGTTTGTCTCGAAGGACAAATACATTTACAGGAGTCACCTCTTGGGGTCTTGGC 2370

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db 2371 TGTGACGCCCAATAAGCCCTGGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2430

Qy 201 GluGlyValMetArgAsnAsn 207
Db 2431 GAGGGAGTGATGAGAAATAAT 2451

```

RESULT 7

```

US-07-854-603-1
; Sequence 1, Application US/07854603
; Patent No. 5637492
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Edwards, Richard M
; APPLICANT: Forman, Joan M
; TITLE OF INVENTION: Activatable fibrinolytic and
; anti-thrombotic proteins

```

```

; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,603
; FILING DATE: 19901207
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2753 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2753
; OTHER INFORMATION: /note= "Fig. 2 Plasminogen cDNA"
; OTHER INFORMATION: "sequence"
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 65..121
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 122..2494
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 65..2494
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 54..55
; OTHER INFORMATION: /note= "Ball site"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2564..2565
; OTHER INFORMATION: /note= "Sphi site"
US-07-854-603-1

Alignment Scores:
Pred. No.: 7.78e-130 Length: 2753
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-992-095B-54 (1-207) x US-07-854-603-1 (1-2753)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1874 ATGCACCTTCTGTGGAGGCACCTTGATATCCCGAGTGGGTGTTGACTGCTGCCACTGC 1933

Qy 21 LeuGluYsSerProArgProSerSerTyRlyValIleLeuGlyAlaHisGluVal 40
Db 1934 TTGGAGAAGTCCCCAAGGCCCTTCTCTACAGGTGCATCTGGGTGGCACCAAGAAGTG 1993

```

QY 41 AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60
DB 1994 AATCTGACCGCATGTTGAGAAATAGAGTGTCTAGGCTGTTCTTGAGCCACACGA 2053
QY 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80
DB 2054 AAGATATTGCTTGTAAAGCTAAGCAGTCTGCGCTCATCACTGACAAAGTAATCCCA 2113
QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
DB 2114 GCTTGTCTGCGCATCCCAATATTATGTGCTGCTGACCGGACCGAATGTTTCATCACTGGC 2173
QY 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
DB 2174 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCCACTCCCTGTG 2233
QY 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
DB 2234 ATTCAGAAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAGAGTCCAATCCACGAA 2293
QY 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyPro 160
DB 2294 CTCTGTCTGCGCATTTGGCCGAGGCACTGACAGTTGCCAGGCTGACAGTGGAGGTCT 2353
QY 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
DB 2354 CTGTTTGTCTCGAAGGACAAATACATTTTACAGAGTCACTTCTTGGGTCTTGGC 2413
QY 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
DB 2414 TGTGACGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2473
QY 201 GluGlyValMetArgAsnAsn 207
DB 2474 GAGGAGGTGATGAGAAATAAT 2494

RESULT 8

US-09-992-095b-54 (1-207) x US-08-643-219-12 (1-2497)
; Sequence 12, Application US/08643219
; Patent No. 5801146
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/643,219
; FILING DATE: 06-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 5940.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; TELEFAX: 847-938-2623
; TELEX:

INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2497 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-643-219-12

Alignment Scores:
Pred. No.: 8,95e-130 Length: 2497
Score: 1112.00 Matches: 206
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 0
Query Match: 99.91% Indels: 0
DB: 1 Gaps: 0

US-09-992-095b-54 (1-207) x US-08-643-219-12 (1-2497)

QY 1 MetHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
DB 1859 ATGCACCTTCTGTGGAGGCACCTTGATATCCCGAGAGTGGGTGTGACTGTGCCCATGTC 1918
QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
DB 1919 TTGGAGAGGTCCCAAGGCCTTCATCTACAGGTCTATCTGCGTGCACACCAAGAAGTG 1978
QY 41 AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60
DB 1979 AATCTCGAACCGCATGTTACAGAAATAGAAAGTGTCTAGGCTGTTCTTGGAGCCACACGA 2038
QY 61 LysAspIleAlaLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
DB 2039 AAGATATTGCTTGTCTAAGCTAAGCAGTCTGCGCTCATCACTGACAAAGTAATCCCA 2098
QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
DB 2099 GCTTGTCTGCCATCCCAATATTATGTTGCTGCTGACCGGACCGAATGTTTCTGCTCCTGGC 2158
QY 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLysGluAlaGlnLeuProVal 120
DB 2159 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCCAAGTCCCTGTG 2218
QY 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
DB 2219 ATTGAGATTAAGTGTGCATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 2278
QY 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyPro 160
DB 2279 CTCTGTCTGCGCATTTGGCCGAGGCACTGACAGTTGCCAGGCTGACAGTGGAGGTCTCT 2338
QY 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
DB 2339 CTGTTTGTCTGAGAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 2398
QY 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
DB 2399 TGTGACGCCCAATTAAGCTGTGTCTAAGTTTCAAGGTTTGTACTTGGATT 2458
QY 201 GluGlyValMetArgAsnAsn 207
DB 2459 GAGGAGGTGATGAGAAATAAT 2479

RESULT 9

US-09-131-995-12
; Sequence 12, Application US/09131995
; Patent No. 5972896
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories
 STREET: 100 Abbott Park Road
 CITY: Abbott Park
 STATE: IL
 COUNTRY: USA
 ZIP: 60064
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/131,995
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/832,087
 FILING DATE: 03-APR-1997
 APPLICATION NUMBER: 08/643,219
 FILING DATE: 06-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Casuto, Dianne
 REGISTRATION NUMBER: 40,943
 REFERENCE/DOCKET NUMBER: 5940.US.P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 847-938-3137
 TELEFAX: 847-938-2623
 TELEX:
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2497 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 US-09-131-995-12

Alignment Scores:
 Pred. No.: 8,95e-130 Length: 2497
 Score: 1112.00 Matches: 206
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 0
 Query Match: 99.91% Indels: 0
 DB: 2 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-131-995-12 (1-2497)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1859 ATGCACCTCTGTGGAGGCACCTTGATATCCCGAGTGGGTGTGACTGTGCCCACTGC 1918
 Qy 21 LeuGluYsSerProArgProSerSerTyrlYsValIleLeuGlyAlaHisGlnGluVal 40
 Db 1919 TTGGAGAAGTCCCCAAGCCCTTATCTTACAGGTATCTCTGGGTGCACCAAGAAGTG 1978
 Qy 41 AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1979 AATCTCGAAGCCGATGTTACAGAAATAGAGTGTAGGCTGTCTTGGAGCCCAACAGA 2038
 Qy 61 LysAspIleAlaLeuLeuYsSerSerProAlaValIleThrAspLysValIlePro 80
 Db 2039 AAAGATATTCCTGTGTAAGCTAAGCAGTCTCGCGTCATCACTGACAAAGTAATCCCA 2098
 Qy 81 AlaCysLeuProSerProAsnTyrrValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 2099 GCITGTCTGCCATCCCCAAATATGTGTGCTGCTGACCGACCGAATGTTTGTGTCACCTGC 2158
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuYsGluAlaGlnLeuProVal 120
 Db 2159 TGGGGAGNAACCAAGGTACTTTTGGAGTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 2218
 Qy 121 IleGluAsnLysValCysAsnArgTyrrGluPheLeuAsnGlyArgValGlnSerThrGlu 140

Db 2219 ATTGAGAAATAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCATCCACCGAA 2278
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2279 CTCTGTCTGGCATTGTGGCCGGAGGCACAGTGTGCCAGGTGACAGTGGAGGTCT 2338
 Qy 161 LeuValCysPheGluLysAspLysTyrrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2339 CTGGTTTGTCTCGAGAAAGGCAATACATTTTACAAGAGTCACTTCTTGGGGTCTTGGC 2398
 Qy 181 CysAlaArgProAsnLysProGlyValTyrrValArgValSerArgPheValThrTrpIle 200
 Db 2399 TGTGCACGCCCAATAAGCTGTGTCTATGTGTCTTCAAGTTTGTACTTGGATT 2458
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db 2459 GAGGGAGTGTAGATAATAAT 2479

RESULT 10

US-08-832-087B-12
 ; Sequence 12, Application US/08832087B
 ; Patent No. 5981484
 ; GENERAL INFORMATION:
 ; APPLICANT: Davidson, Donald J.
 ; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
 ; TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Abbott Laboratories
 ; STREET: 100 Abbott Park Road
 ; CITY: Abbott Park
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60064
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/832,087B
 ; FILING DATE: 03-APR-1997
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/643,219
 ; FILING DATE: 06-MAY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Casuto, Dianne
 ; REGISTRATION NUMBER: 40,943
 ; REFERENCE/DOCKET NUMBER: 5940.US.P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 847-938-3137
 ; TELEFAX: 847-938-2623
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2497 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna
 ; US-08-832-087B-12

Alignment Scores:
 Pred. No.: 8,95e-130 Length: 2497
 Score: 1112.00 Matches: 206
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 0
 Query Match: 99.91% Indels: 0
 DB: 2 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-832-087B-12 (1-2497)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1859 ATGCACCTCTGTGGAGGACCTTGTATATCCACAGTGGGTGTGACTGTGCTGCCACTGC 1918
 Qy 21 LeuGluLysSerProArgProSerSerTyrlsValleLeuGlyAlaHisGlnGluVal 40
 Db 1919 TTGGAGAAGTCCCAAGGCTTCTATCTACAGGTCTCTGGGTGCACCAAGAAGTG 1978
 Qy 41 AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60
 Db 1979 AATCTCGAACCGCATTTGAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 2038
 Qy 61 LysAspIleAlaLeuLeuLysSerSerProAlaValIleThrAspLysValIlePro 80
 Db 2039 AAAGATATTGCTTGTAAAGCTAAGCAGTCTCTGCCGTCTCATCTGACAAAGTAAATCCCA 2098
 Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
 Db 2099 GCTTGTCTGCATCCCAATATATGTGTCTGCTGACCGGACCGAATGTTTCGTCACTGGC 2158
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 2159 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCCAAGTCCCTGTG 2218
 Qy 121 IleGluAsnLysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 2219 ATTGAGATAAAGTGTGCATCGTATGAGTTCCTGAATGGAAGAGTCCATCCACCGAA 2278
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2279 CTCTGTCTGGCATTTGGCCGAGGACACTGACAGTTGCCAGGTGCACAGTGGAGGTCT 2338
 Qy 161 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2339 CTGCTTGTCTCGAAGAGCAAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 2398
 Qy 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
 Db 2399 TGTGACGCGCCCAATAGCTGGTGTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2458
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db 2459 GAGGAGTGATGAGAAATAAT 2479

RESULT 11

US-08-851-350-12
 ; Sequence 12, Application US/08851350
 ; Patent No. 6057122

GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories
 ; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES, AND METHODS
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME, AND METHODS
 ; TITLE OF INVENTION: FOR INHIBITING ANGIOGENESIS

NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories

; STREET: 100 Abbott Park Road

; CITY: Abbott Park

; STATE: IL

; COUNTRY: USA

; ZIP: 60064-3500

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/851,350

; FILING DATE: 05-MAY-1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Casuto, Dianne
 ; REGISTRATION NUMBER: 40,943
 ; REFERENCE/DOCKET NUMBER: 5940.US.P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 847-938-3137
 ; TELEFAX: 847-938-2623
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2497 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-851-350-12

Alignment Scores: 8,95e-130 Length: 2497
 Pred. No.: 1112.00 Matches: 206
 Score: 100.00% Conservative: 1
 Percent Similarity: 100.00%
 Best Local Similarity: 99.52% Mismatches: 0
 Query Match: 99.91% Indels: 0
 DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-851-350-12 (1-2497)

Qy 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
 Db 1859 ATGCACCTCTGTGGAGGACCTTGTATATCCACAGTGGGTGTGACTGTGCTGCCACTGC 1918
 Qy 21 LeuGluLysSerProArgProSerSerTyrlsValleLeuGlyAlaHisGlnGluVal 40
 Db 1919 TTGGAGAGTGTCCCAAGGCTTCTATCTACAGGTCTCTTGGGTGCACCAAGAAGTG 1978
 Qy 41 AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60
 Db 1979 AATCTCGAACCGCATTTGAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 2038
 Qy 61 LysAspIleAlaLeuLeuLysSerSerProAlaValIleThrAspLysValIlePro 80
 Db 2039 AAAGATATTGCTTGTAAAGCTAAGCAGTCTCTGCCGTCTCATCTGACAAAGTAAATCCCA 2098
 Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
 Db 2099 GCTTGTCTGCATCCCAATATATGTGTCTGCTGACCGGACCGAATGTTTCGTCACTGGC 2158
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 2159 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCCAAGTCCCTGTG 2218
 Qy 121 IleGluAsnLysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 2219 ATTGAGATAAAGTGTGCAATCGTATGAGTTCCTGAATGGAAGAGTCCATCCACCGAA 2278
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 2279 CTCTGTCTGGCATTTGGCCGAGGACACTGACAGTTGCCAGGTGCACAGTGGAGGTCT 2338
 Qy 161 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 2339 CTGCTTGTCTCGAAGAGCAAAATAATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 2398
 Qy 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
 Db 2399 TGTGACGCGCCCAATAGCTGGTGTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2458
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db 2459 GAGGAGTGATGAGAAATAAT 2479

RESULT 12

US-09-132-154-12

; Sequence 12, Application US/09132154

Patent No. 6251867
GENERAL INFORMATION:
APPLICANT: Davidson, Donald J.
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/132,154
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/832,087
FILING DATE: 03-APR-1997
APPLICATION NUMBER: 08/643,219
FILING DATE: 06-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 5940 US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2497 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-132-154-12

Alignment Scores:
Pred. No.: 8.95e-130 Length: 2497
Score: 1112.00 Matches: 206
Percent Similarity: 100.00% Conservativeness: 1
Best Local Similarity: 99.52% Mismatches: 0
Query Match: 99.91% Indels: 0
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-132-154-12 (1-2497)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1859 ATGCACCTCTCTGAGGACACCTTGATATCCACAGTGGGTGTTGACTGTGCTGCCACTGC 1918
Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Db 1919 TTGGAGAAGTCCCCAAGCCCTTCATCTCTACAGGTGTCATCTGGGTGCACACCAAGAAGTG 1978
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1979 AATCTCGAACCCGATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACAGA 2038
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 2039 AAAGATATTGCTTGTAAAGCTAAGCAGTCTCTGCCGTGTCATCTGACAAAGTAATCCCA 2098
Qy 81 AlaCysLeuProSerProAsnTyrlValIleAlaSerArgThrGluCysPheIleThrGly 100
Db 2099 GCTTGTCTGCATCCCCAAATATTATGTGTGCTGACCGGACCGAATGTTTGTCTACCTGGC 2158

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 2159 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCAGCTCCCTGTG 2218
Qy 121 IleGluAsnLysValCysAsnArgTyrlGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 2219 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGNA 2278
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2279 CTCTGTGCTGGCATTTTGGCCGGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCT 2338
Qy 161 LeuValCysPheGluLysAspLysTyrlLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2339 CTGGTTTGTCTCGAAGAGGCAAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 2398
Qy 181 CysAlaArgProAsnLysProGlyValTyrlValArgValSerArgPheValThrTrpIle 200
Db 2399 TGTGCACGCCCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2458
Qy 201 GluGlyValMetArgAsnAsn 207
Db 2459 GAGGAGCTGATGAGAAATAAT 2479

RESULT 13
US-08-379-621-1
Sequence 1, Application US/08379621
Patent No. 5645833
GENERAL INFORMATION:
APPLICANT: DAWSON, Keith
APPLICANT: GILBERT, Richard James
TITLE OF INVENTION: INHIBITOR RESISTANT SERINE PROTEASES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/379,621
FILING DATE: 03-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/01632
FILING DATE: 03-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Hollie L.
REGISTRATION NUMBER: 31,321
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 942-8400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 690 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 1..690
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start=1

```
; OTHER INFORMATION: /function= "encodes plasmin protease domain"
; OTHER INFORMATION: /product= "nucleotide with corresponding
; OTHER INFORMATION: /protein=
; OTHER INFORMATION: /number= 1
US-08-379-621-1
Alignment Scores:
Pred. No.: 7.04e-130 Length: 690
Score: 1106.00 Matches: 206
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.37% Indels: 0
DB: 1 Gaps: 0
US-09-992-095B-54 (1-207) x US-08-379-621-1 (1-690)
Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 70 ATGCACCTCTGTGGAGGACCTTGATATCCCGAGTGGGTGTGACTGTGCCCCACTGC 129
Qy 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db 130 TTGGAGAAGTCCCCAAGCCCTTCTCTACAGGTCTCTGGGTGCACACCAAGAAGTG 189
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 190 AATCTCGAACCGCATGTGTACAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 249
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 250 AAAGATATTGCTTGTCTAAAGCTTAAGCAGTCTCTGCCGTCTACCTGACCAAGTAATCCCA 309
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 310 GCTTGTCTGCCATCCCCCAAAATATGTGTCTGCTACCGACCGAATGTTTCTACCTGCG 369
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 370 TGGGAGAAACCCCAAGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCAGCTCCCTGTG 429
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 430 ATTGAGAATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 489
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 490 CTCTGTCTGGGCAATTTGGCCGAGGCACTGACAGTTGCCAGGTGACAGTGGAGGTCT 549
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 550 CTGGTTTGTCTCGAGAGGACCAATACATTTTACAGGAGTCACTTCTTGGGTCTTGGC 609
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 610 TGTGACGCCCCAATAAGCTGTGTCTATGTTCTGTGTTTCAAGGTTTGTACTTGGATT 669
Qy 201 GluGlyValMetArgAsn 207
Db 670 GAGGAGTGTGAGAAATAAT 690
RESULT 14
US-08-889-078-1
; Sequence 1, Application US/08889078
; Patent No. 5932213
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith Martyn
; APPLICANT: Richard James Gilbert
; TITLE OF INVENTION: MODIFIED PLASMIN PRECURSORS WITH RESISTANCE TO
; TITLE OF INVENTION: INHIBITORS OF PLASMIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
```



```
Db 370 TGGGAGAAACCCAGGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCTGTG 429
Qy 121 IleGluAenLysValCysAsnArgTyrGluPheLeuAenGlyValGlnSerThrGlu 140
Db 430 ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 489
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyValSerThrGlyPro 160
Db 490 CTCTGTGTGGGCATTTGGCCGAGGACACTGACAGTTCAGAGGTGACAGTGGAGTCTT 549
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerThrGlyLeuGly 180
Db 550 CTGGTTTCTTCGAGAAGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 609
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db 610 TGTGCACGCCCCCAATAGCCTGGTGTCTATGTTGTGTTCAGGTTTGTACTTGGATT 669
Qy 201 GluGlyValMetArgAsnAsn 207
Db 670 GAGGAGTGTAGAGAATAAT 690

RESULT 15
US-09-820-002-1
; Sequence 1, Application US/09820002
; Patent No. 6482630
; GENERAL INFORMATION:
; APPLICANT: Yan, Weinuu
; APPLICANT: Ye, Jane
; APPLICANT: DiFrancesco, Valentina
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO01194
; CURRENT APPLICATION NUMBER: US/09/820,002
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: HUMAN
US-09-820-002-1

Alignment Scores:
Pred. NO.: 6.55e-46 Length: 1615
Score: 444.00 Matches: 92
Percent Similarity: 59.63% Conservative: 38
Best Local Similarity: 42.20% Mismatches: 66
Query Match: 39.89% Indels: 22
DB: 4 Gaps: 7

US-09-992-095B-54 (1-207) x US-09-820-002-1 (1-1615)
Qy 2 HisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCysLeu 21
Db 610 CACCTCTGTGGGGATCCCTGCTCTCCGGGACCTGGGTGTCGACAGCCGCCCACTGCTTC 669
Qy 22 ---GluLysSerProArgProSerSerTyrLysValIleLeuGlyValAlaHisGlnGluVal 40
Db 670 CCGAGCGGAACCGGGTCTCTCCGATGGCGAGTGTTCGGGTGCGCGTGGCCGAGGCC 729
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeu----- 54
Db 730 TCT-----CCCCACGGTCTGAGTGGGGTGGCGGTGAGGTGCTACACGGGGCTAT 783
Qy 55 -----PheLeuGluProThrArgLys-----AspIleAlaLeuLysLeu 68
Db 784 CTTCCCTTTTCGGGACCCCAACAGCGAGGAGAACAGCATATTTGCCCTGTTCCACCTC 843
Qy 69 SerSerProAlaValIleThrAspLysValIleProAlaCysLeuProSerProAsnTyr 88
```

```
Db 844 TCCAGTCCCTCCCTCCCTCCACAGAAATACATCCAGGCTGTGTGCCTCCACAGCTCCGCGCCAG 903
Qy 89 ValValAlaAapArgThrGluCysPheIleThrGlyTyrGlyGluThrGlyGlyThrPhe 108
Db 904 CCCCTGTGGTGGATGGCAAGATCTGTACCGTGACGGGCTGGGGCAACACCGCAG---TACTAT 960
Qy 109 Gly-----AlaGlyLeuLeuLysGluAlaGlnLeuProValIleGluAsnLysValCys 126
Db 961 GGCCAAACAGGCGCGGGTACTCCAGAGGCTCGAGTCCCCCAATAATCAGCAATGATGTCTGTC 1020
Qy 127 AsnArgTyrGluPheLeuAenGlyArgValGlnSerThrGluLeuCysAlaGlyHisLeu 146
Db 1021 AATGGCGTGACTTCTTATGGAAACACAGATCAAGCCCAAGATGTTCTGTGCTGGCTACCCC 1080
Qy 147 AlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLys 166
Db 1081 GAGGGTGGCATTGATGCTGCGAGGGCGACAGCGGTGTCCTTTGTGTGTGAGGACAGC 1140
Qy 167 -----AspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGlyCysAla 182
Db 1141 ATCTCTCGGACGCCACCGTTGGCGGCTGTGTGCATTGTGAGTTGGGGCACTGGCTGTGCC 1200
Qy 183 ArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db 1201 CTGGCCCAAGAGCCAGCGCTCTACACCAAGTCACTGCTGCTTCCGGGAGTGGATC 1254
```

Search completed: September 14, 2004, 01:19:42
Job time : 91 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 14, 2004, 00:35:23 ; Search time 466 Seconds
(without alignments)
2235.033 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 1113
Sequence: 1 MHFCGGLISPEWVLAHA.....GVYVRVSRFTWIEGVMRNN 207

Scoring table:

BLOSUM62			
Xgapop 10.0 , Xgapext 0.5			
Ygapop 10.0 , Ygapext 0.5			
Fgapop 6.0 , Fgapext 7.0			
Delop 6.0 , Delext 7.0			

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US9992095/runat_10092004.104840.5038/app_query.fasta_1.391
-DB=PublishedApplications_NA -QFMT=factap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -MAP=US9992095 @cgn 1 1 520 @runat_10092004.104840.5038
-NCPU=6 -ICPU=3 -NO_WARL -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-10-450-976-3

; Sequence 3, Application US/10450976

; Publication No. US20040071676A1

; GENERAL INFORMATION:

; APPLICANT: COLLEN, Desire Jose

; APPLICANT: NAGAI, Nubuo

; APPLICANT: LAROCHE, Yves

; TITLE OF INVENTION: A Yeast Expression Vector and a Method

; TITLE OF INVENTION: of Making a Recombinant Protein by Expression in a Yeast

; TITLE OF INVENTION: Cell

; FILE REFERENCE: 50304/005001

; CURRENT APPLICATION NUMBER: US/10450.976

; CURRENT FILING DATE: 2003-06-18

; PRIOR APPLICATION NUMBER: PCT/BE01/00217

; PRIOR FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: GB 0116702

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: GB 0116690

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: GB 0031196

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 10

Sequence 3, Appli
Sequence 53, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 12, Appl
Sequence 125, App
Sequence 125, App
Sequence 125, App
Sequence 3, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 69, Appl
Sequence 69, Appl
Sequence 717, App
Sequence 717, App
Sequence 771, App
Sequence 820, App
Sequence 858, App
Sequence 155, App
Sequence 59, Appl
Sequence 11, Appl
Sequence 68, Appl
Sequence 68, Appl
Sequence 68, Appl

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(750)
US-10-450-976-3

Alignment Scores:
Pred. No.: 2,99e-139 Length: 750
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-450-976-3 (1-750)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 127 ATGCACCTTCCTGTGAGGACCCCTTGATATCCAGAGTGGGTGTGACTGTGCCCACTGC 186

Qy 21 LeuGluLysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Db 187 TTGGAGAAGTCCCAAGGCCCTTCATCTACAGGTCTATCTGGGTGCACACCAAGAAGTG 246

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 247 AATCTCGAACCGCATGTTCCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCCAACGA 306

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 307 AAAGATATATGCTTGTCTAAAGCTAAGCAGTCTCCGCCGTCTACCTGACAAAGTAATCCCA 366

Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
Db 367 GCTTGTCTGCCATCTCCCAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 427 TGGGAGAGAAACCAAGGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCCAAGTCCCTGTG 486

Qy 121 IleGluAsnLysValCysAsnArgTyrlsPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 487 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTCTGAATGGAAGAGTCCCAATCCACCGAA 546

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 547 CTCTGTGCTGGGCATTTGGCCGGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCTCT 606

Qy 161 LeuValCysPheGluLysAspLysTyrlsIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 607 CTGGTTTGTCTCGAAGAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGCTTTGGC 666

Qy 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
Db 667 TGTGACGCCCAATAAGCTGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 726

Qy 201 GluclyValMetArgAsnAsn 207
Db 727 GAGGAGTGTATGAGAAATAT 747

RESULT 2
US-10-450-976-5
; Sequence 5, Application US/10450976
; Publication No. US20040071676A1
; GENERAL INFORMATION:
; APPLICANT: COLLEN, Desire Jose
; APPLICANT: NAGAI, Nubuo
; APPLICANT: LAROCHE, Yves
; TITLE OF INVENTION: A Yeast Expression Vector and a Method

; TITLE OF INVENTION: of Making a Recombinant Protein by Expression in a Yeast
; ID NO INVENTION: Cell
; FILE REFERENCE: 50304/005001
; CURRENT APPLICATION NUMBER: US/10/450,976
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/BE01/00217
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: GB 0116702
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0116690
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0031196
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1047)
US-10-450-976-5

Alignment Scores:
Pred. No.: 4,83e-139 Length: 1047
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-450-976-5 (1-1047)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 424 ATGCACCTTCCTGTGAGGACCCCTTGATATCCAGAGTGGGTGTGACTGTGCCCACTGC 483

Qy 21 LeuGluLysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Db 484 TTGGAGAAGTCCCAAGGCCCTTTCATCTCAAGGTCTATCTGCTGCTGCTGCTGCTGCTGCTG 543

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 544 AATCTCGAACCGCATGTTCCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCCAACGA 603

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 604 AAAGATATATGCTTGTCTAAAGCTAAGCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 663

Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
Db 664 GCTTGTCTGCCATCTCCCAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 724 TGGGAGAGAAACCAAGGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCCAAGTCCCTGTG 783

Qy 121 IleGluAsnLysValCysAsnArgTyrlsPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 784 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTCTGAATGGAAGAGTCCCAATCCACCGAA 843

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 844 CTCTGTGCTGGGCATTTGGCCGGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGTCTCT 903

Qy 161 LeuValCysPheGluLysAspLysTyrlsIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 904 CTGGTTTGTCTCGAAGAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGCTTTGGC 963

Qy 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
Db 964 TGTGACGCCCAATAAGCTGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1023
```

Qy 201 GluGlyValMetArgAsnAsn 207
 Db 1024 GAGGAGTGATGAGAAATAAT 1044

RESULT 3

US-09-992-600A-53
 ; Sequence 53, Application US/09992600A
 ; Publication No. US20030027161A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benjanin, Stephane
 ; APPLICANT: Tanaka, Hiroaki
 ; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 91. US 4, DIV
 ; CURRENT APPLICATION NUMBER: US/09/992.600A
 ; CURRENT FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 09/924,340
 ; PRIOR FILING DATE: 2001-08-06
 ; PRIOR APPLICATION NUMBER: PCT/IB01/01715
 ; PRIOR FILING DATE: 2001-08-06
 ; PRIOR APPLICATION NUMBER: US 60/305,456
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 60/302,277
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/298,698
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 60/293,574
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: JPatent
 ; SEQ ID NO 53
 ; LENGTH: 1907
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: 1..1043
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1044..1664
 ; FEATURE:
 ; NAME/KEY: 3'UTR
 ; LOCATION: 1665..1907
 ; FEATURE:
 ; NAME/KEY: polyA signal
 ; LOCATION: 1869..1874
 ; FEATURE:
 ; NAME/KEY: polyA site
 ; LOCATION: 1892..1907
 ; US-09-992-600A-53

Alignment Scores:
 Pred. No.: 1.14e-138 Length: 1907
 Score: 1113.00 Matches: 207
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-992-600A-53 (1-1907)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20
 Db 1044 ATGCACCTCTGTGGAGGACCTTCATATCCAGAGTGGGTGTGACTGCTGCCACTGC 1103
 Qy 21 LeuGluYsserProArgProSerSerTyriysValIleLeuGlyAlaHisGlnGluVal 40
 Db 1104 TTGGAGAAGTCCCCAAGGCTTCATCTCTACAAGGTCTCTGGGTGCACACCAAGAAGTG 1163
 Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
 Db 1164 AATCTCGAACCGCATGTTTCAGGAATAAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 1223

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
 Db 1224 AAAGATATTGCTTGTCTAAAGCTTAAGCAGTCTCTCCCGTCTCATCTGACAAAGTAAATCCCA 1283
 Qy 81 AlaCysLeuProSerProAsnTyriysValValAlaAspArgThrGluCysPheIleThrGly 100
 Db 1284 GCTTGTCTGCATATCCCAATATTATGTGTGCTGCTGACCGGACCGAATGTTTCTACTCTGGC 1343
 Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
 Db 1344 TGGGAGAAACCCCAAGGTACTTTTGGAGTGGCCTTCTCAAGGAAGCCACGCTCCCTGTG 1403
 Qy 121 IleGluAsnLysValCysAsnArgTyriysGluPheLeuAsnGlyArgValGlnSerThrGlu 140
 Db 1404 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 1463
 Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
 Db 1464 CTCTGTCTGGGCATTTGGCCGGAGGCATGACAGTTGCCAGGTGACAGTGGAGTCTCT 1523
 Qy 161 LeuValCysPheGluLysAspLysTyriysIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
 Db 1524 CTGCTTGTCTCGAGAGGACAAATACATTTTCAAGGAGTCACTTCTTGGGGTCTTGGC 1583
 Qy 181 CysAlaArgProAsnLysProGlyValTyriysValArgValSerArgPheValThrTrpIle 200
 Db 1584 TGTGCAGCCCAATAGCCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTCTACTTGGATT 1643
 Qy 201 GluGlyValMetArgAsnAsn 207
 Db 1644 GAGGAGTGATGAGAAATAAT 1664

RESULT 4

US-09-924-340-53
 ; Sequence 53, Application US/09924340
 ; Publication No. US20030027248A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benjanin, Stephane
 ; APPLICANT: Tanaka, Hiroaki
 ; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 91. US 2, REG
 ; CURRENT APPLICATION NUMBER: US/09/924,340
 ; CURRENT FILING DATE: 2001-08-06
 ; PRIOR APPLICATION NUMBER: US 60/305,456
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 60/302,277
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/298,698
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 60/293,574
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 112
 ; SOFTWARE: JPatent
 ; SEQ ID NO 53
 ; LENGTH: 1907
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: 1..1043
 ; NAME/KEY: CDS
 ; LOCATION: 1044..1664
 ; NAME/KEY: 3'UTR
 ; LOCATION: 1665..1907
 ; NAME/KEY: polyA signal
 ; LOCATION: 1869..1874
 ; NAME/KEY: polyA site
 ; LOCATION: 1892..1907
 ; US-09-924-340-53

Alignment Scores:
 Pred. No.: 1.14e-138 Length: 1907
 Score: 1113.00 Matches: 207

Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-924-340-53 (1-1907)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db ATGCACCTCTGTGGAGGACCTTGTATATCCAGAGTGGGTGTACTGTGCCCATGTC 1103

Qy 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db TTGGAGAAGTCCCAAGGCTTCATCTACAGGTCTCTGGGTGCACACCAAGAAGTG 1163

Qy 41 AsnLeuGluProHisValGlnGluIleValSerArgLeuPheLeuGluProThrArg 60
Db AATCTCGAACCGCATGTCAGGAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 1223

Qy 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80
Db AAAGATATTGCTTGTCTAAGCTAAGCAGTCTCTGCCGTCACTGACAAAGTAAATCCCA 1283

Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db GCTTGTCTGCCATCCCAATATTATGTGTGCTGACCGGACCGAATGTTTCATCACTGGC 1343

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCCAAGTCCCTGTG 1403

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCATCCACCGAA 1463

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db CTCTGTGCTGGCATTTGGCCGAGGACCTGACAGTTGCCAGGTGTGACAGTGGAGGTCT 1523

Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db CTGGTTTGTCTCGAAGAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 1583

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db GTGCACGCCCAATAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1643

Qy 201 GluGlyValMetArgAsnAsn 207
Db GAGGAGTGTAGAGAAATAAT 1664

RESULT 5

US-09-992-095B-53
; Sequence 53, Application US/09992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US.DIV
; CURRENT APPLICATION NUMBER: US/09/992,095B
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574

PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: Jpatent
SEQ ID NO 53
LENGTH: 1907
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..1043
FEATURE:
NAME/KEY: CDS
LOCATION: 1044..1664
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1665..1907
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 1869..1874
FEATURE:
NAME/KEY: polyA_site
LOCATION: 1892..1907
US-09-992-095B-53

Alignment Scores:
Pred. No.: 114e-138 Length: 1907
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-992-095B-53 (1-1907)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1044 ATGCACCTCTGTGGAGGACCTTGTATATCCAGAGTGGGTGTACTGTGCCCATGTC 1103

Qy 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db 1104 TTGGAGAAGTCCCAAGGCTTCTTCAAGGTCTATCTCTGGGTGCACACCAAGAAGTG 1163

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1164 AATCTCGAACCGCATGTCAGGAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 1223

Qy 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80
Db 1224 AAAGATATTGCTTGTCTAAGCTAAGCAGTCTCTGCCGTCACTGACAAAGTAAATCCCA 1283

Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 1284 GCTTGTCTGCCATCCCAATATTATGTTGCTGACCGGACCGAATGTTTCATCACTGGC 1343

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 1344 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCCAAGTCCCTGTG 1403

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1404 ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCATCCACCGAA 1463

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1464 CTCTGTGCTGGCATTTGGCCGAGGACCTGACAGTTGCCAGGTGTGACAGTGGAGGTCT 1523

Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1524 CTGGTTTGTCTCGAAGAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 1583

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 1584 GTGCACGCCCAATAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1643

```
Qy 201 GluGlyValMetArgAsnAsn 207
Db 1644 GAGGAGTGATGAGAAATAAT 1664

RESULT 6
US-09-999-570-53
; Sequence 53, Application US/09999570
; Publication No. US20030170628A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: G-0910US08DIV
; CURRENT APPLICATION NUMBER: US/09/999,570
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 53
; LENGTH: 1907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..1043
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1044..1664
; NAME/KEY: polyA signal
; LOCATION: 1869..1874
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 1892..1907
US-09-999-570-53

Alignment Scores:
Pred. No.: 1.14e-138 Length: 1907
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-992-095B-54 (1-207) x US-09-999-570-53 (1-1907)

Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1044 ATGCACCTCTGTGGAGGACCTTCATATCCCGAGTGGGTGTGACTGCTGCCACTGC 1103

Qy 21 LeuGluIysSerProArgProSerSerTyrIysValIleLeuGlyAlaHisGlnGluVal 40
Db 1104 TTGGAGAAGTCCCCAAGCCCTTCATCTACAAAGTCTCCTGGGTGCACACCAAGAAGTG 1163

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1164 AATCTCGAACCGCATGTTTCAGGAAATAGAAAGTGTCTAGGCTGTCTTGGAGCCACACGA 1223

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspIysValIlePro 80
Db 1224 AAAGATATTGCTTAAAGCTAAGCAGTCTCTCCGCTCATCTACATGACAAAGTAATCCCA 1283

Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAlaSerArgThrGluCysPheIleThrGly 100
Db 1284 GCITGTCTGCCATCCCCAAATTTATGTGTCTGCTGACCGGACCGAATGTTTCTACTCTGC 1343

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 1344 TGGGAGAAACCCCAAGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCAGCTCCCTCTGTG 1403

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1404 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 1463

Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1464 CTCTGTCTGGGCATTTTGGCGGAGGACCTGACAGTTGCCAGGTTGACAGTGGAGGTCTT 1523

Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1524 CTGTTTGTTCGAGAAGGACAAATACATTTTACAGGAGTCACTTCTTGGGCTCTTGGC 1583

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 1584 TGTGACGCCCAATAAGCTGTGTCTATGTTCTGTGTTTCAAGGTTTGTACTTGGATT 1643

Qy 201 GluGlyValMetArgAsnAsn 207
Db 1644 GAGGAGTGATGAGAAATAAT 1664

RESULT 7
US-10-000-489-53
; Sequence 53, Application US/10000489
; Publication No. US20030092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91 US6 DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 53
; LENGTH: 1907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..1043
; NAME/KEY: CDS
; LOCATION: 1044..1664
; NAME/KEY: polyA signal
; LOCATION: 1665..1907
; NAME/KEY: polyA site
; LOCATION: 1869..1874
; NAME/KEY: polyA site
; LOCATION: 1892..1907
US-10-000-489-53
```

Alignment Scores:
Pred. No.: 1.14e-138 Length: 1907
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-000-489-53 (1-1907)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1044 ATGCACCTCTGTGGAGCACCCTTGATATCCAGAGTGGGTGTTGACTGCTGCCCACTGC 1103

Qy 21 LeuGluIysSerProArgProSerSerTyriysValIleLeuGlyAlaHisGlnGluVal 40
Db 1104 TTGGAGAAGTCCCAAGCCCTTCATCTACAAAGTCTATCTGGGTGCACCAAGAAGTG 1163

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1164 AATCTCGAACCGCATGTTACAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACGA 1223

Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 1224 AAAGATATTGCTTGTAAAGCTAAGCAGTCTCTGCCGTCACTGACAAAGTAATCCCA 1283

Qy 81 AlaCysLeuProSerProAsnTyriysValAlaAspArgThrGluCysPheIleThrGly 100
Db 1284 GCTTGTCTGCCATCCCAAGCTTCTTGGAGCTGGCTTCTCAAGGAAGCCCACTCCCTGTG 1343

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuIysGluAlaGlnLeuProVal 120
Db 1344 TGGGAGAAACCCAGGTAATTTTGGAGCTGGCTTCTCAAGGAAGCCCACTCCCTGTG 1403

Qy 121 IleGluAsnLysValCysAsnArgTyriysValIleLeuGlnSerThrGlu 140
Db 1404 ATTGAGAATAAAGTGTGCATCGTATGAGTTTCTGAATGGAAGAGTCCATCCACCGAA 1463

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1464 CTCTGTCTGGCATTTGGCCGAGGCACTGACAGTTGCCAGGTGACAGTGGAGGTCTT 1523

Qy 161 LeuValCysPheGluIysAspLysTyriysValIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1524 CTGTTGCTTTCGAGAAGGACAAATACATTTTCAAGAGGTCATCTCTTGGGGTCTTGGC 1583

Qy 181 CysAlaArgProAsnLysProGlyValTyriysValArgValSerArgPheValThrTriple 200
Db 1584 TGTGCAGCCCAATAGCCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1643

Qy 201 GluGlyValMetArgAsnAsn 207
Db 1644 GAGGAGTGTATGAGAAATAAT 1664

RESULT 8

US-10-000-986-53
; Sequence 53, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US.DIV
; CURRENT APPLICATION NUMBER: US/10/000.986
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277

US-09-992-095B-54 (1-207) x US-10-000-986-53 (1-1907)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1044 ATGCACCTCTGTGGAGCACCCTTGATATCCAGAGTGGGTGTTGACTGCTGCCCACTGC 1103

Qy 21 LeuGluIysSerProArgProSerSerTyriysValIleLeuGlyAlaHisGlnGluVal 40
Db 1104 TTGGAGAAGTCCCAAGCCCTTCATCTACAAAGTCTATCTGGGTGCACCAAGAAGTG 1163

Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1164 AATCTCGAACCGCATGTTACAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACGA 1223

Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 1224 AAAGATATTGCTTGTAAAGCTAAGCAGTCTCTGCCGTCACTGACAAAGTAATCCCA 1283

Qy 81 AlaCysLeuProSerProAsnTyriysValAlaAspArgThrGluCysPheIleThrGly 100
Db 1284 GCTTGTCTGCCATCCCAAGCTTCTTGGAGCTGGCTTCTCAAGGAAGCCCACTCCCTGTG 1343

Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuIysGluAlaGlnLeuProVal 120
Db 1344 TGGGAGAAACCCAGGTAATTTTGGAGCTGGCTTCTCAAGGAAGCCCACTCCCTGTG 1403

Qy 121 IleGluAsnLysValCysAsnArgTyriysValIleLeuGlnSerThrGlu 140
Db 1404 ATTGAGAATAAAGTGTGCATCGTATGAGTTTCTGAATGGAAGAGTCCATCCACCGAA 1463

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1464 CTCTGTCTGGCATTTGGCCGAGGCACTGACAGTTGCCAGGTGACAGTGGAGGTCTT 1523

Qy 161 LeuValCysPheGluIysAspLysTyriysValIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1524 CTGTTGCTTTCGAGAAGGACAAATACATTTTCAAGAGGTCATCTCTTGGGGTCTTGGC 1583

Qy 181 CysAlaArgProAsnLysProGlyValTyriysValArgValSerArgPheValThrTriple 200
Db 1584 TGTGCAGCCCAATAGCCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1643

Qy 201 GluGlyValMetArgAsnAsn 207
Db 1644 GAGGAGTGTATGAGAAATAAT 1664


```
Oy 201 GluGlyValMetArgAsnAsn 207
Db 1644 GAGGAGTGTGAGAAATAAT 1664

RESULT 9
US-10-154-678-53
; Sequence 53, Application US/10154678
; Publication No. US20030162186A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182.US1.REG
; CURRENT APPLICATION NUMBER: US/10/154,678
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 53
; LENGTH: 1907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..1043
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1044..1664
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 1665..1907
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1869..1874
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 1892..1907
US-10-154-678-53

Alignment Scores:
Pred. No.: 1,14e-138 Length: 1907
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-154-678-53 (1-1907)

Oy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1044 ATGCACCTTCTGTGGAGGCACCTTGATATCCACAGAGTGGGTGTTGACTGTGCTGCCACTGC 1103

Oy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGluVal 40
Db 1104 TTGGAGAGTCCCCAAGGCCCTTCATCTACAGAGTTCATCTGGGTGCACACCAAGAGTG 1163

Oy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1164 AATCTCGNACCGCATGTTCCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACAGA 1223

Oy 61 LysAspIleAlaLeuLeuIysLeuSerProAlaValIleThrAspLysValIlePro 80
Db 1164 AATCTCGNACCGCATGTTCCAGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACAGA 1223
```

```

US-09-992-095B-54 (1-207) x US-10-001-142-53 (1-1907)
Qy 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1044 ATGCACCTTCTGTGAGGACCTTGATATCCCAAGAGTGGGTGTGACTGCTGCCCACTGC 1103
Qy 21 LeuGluIysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db 1104 TTGGAGAGTCCCAAGGCTTCATCTACAGGTCACTCGGTGCACACCAAGAAGTG 1163
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1164 AATCTCGAACCGCATGTTCAAGAAATAGAAAGTGTCTAGGCTGTCTTGGAGCCACACGA 1223
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 1224 AAAGATATTGCTTGTAAAGCTAAGCAGTCTGCGCTCATCACTGACAAAGTAATCCCA 1283
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 1284 GCTTGCTGCATCCCAAAATATGCTGCTGACCGGACCGAATGTTTCATCACTGGC 1343
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 1344 TGGGAGAAACCCCAAGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCACGCTCCCTGTG 1403
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 1404 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTCTGAATGGAAGAGTCAATCCACCGAA 1463
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 1464 CTCTGTGCTGGGCAATTTGGCCGAGGCACTGACAGTTGCCAGGTGCAGTGGAGTCTTGGC 1523
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 1524 CTGTTGCTTCGAGAGGACAAATACATTTTCAAGAGTCACTTCTGGGGTCTTGGC 1583
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 1584 TGTGACGCCCAATAGCCCTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 1643
Qy 201 GluGlyValMetArgAsnAsn 207
Db 1644 GAGGAGTGTATGAGAAATAAT 1664

RESULT 11
US-10-450-976-1
; Sequence 1, Application US/10450976
; Publication No. US20040071676A1
; GENERAL INFORMATION:
; APPLICANT: COLLEN, Desire Jose
; APPLICANT: NAGAI, Nubuo
; APPLICANT: LAROCHE, Yves
; TITLE OF INVENTION: A Yeast Expression Vector and a Method
; TITLE OF INVENTION: Of Making a Recombinant Protein by Expression in a Yeast
; TITLE OF INVENTION: Cell
; FILE REFERENCE: 50304/005001
; CURRENT APPLICATION NUMBER: US/10/450,976
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/BE01/00217
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: GB 0116702
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0116690
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: GB 0031196
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2433

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2433)
US-10-450-976-1

Alignment Scores:
Pred. No.: 1,62e-138 Length: 2433
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-450-976-1 (1-2433)
Qy 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1810 ATGCACCTTCTGTGAGGACCTTGATATCCCAAGAGTGGGTGTGACTGCTGCCCACTGC 1869
Qy 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
Db 1870 TTGGAGAGTCCCAAGGCTTCTATCTACAGGTCACTCTCGGTGCACACCAAGAAGTG 1929
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 1930 AATCTCGAACCGCATGTTCAAGAAATAGAAAGTGTCTAGGCTGTCTTGGAGCCACACGA 1989
Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 1990 AAAGATATTGCTTGTAAAGCTAAGCAGTCTGCGCTCATCACTGACAAAGTAATCCCA 2049
Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 2050 GCTTGCTGCATCCCAAAATATGCTGCTGACCGGACCGAATGTTTCATCACTGGC 2109
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 2110 TGGGAGAAACCCCAAGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCACGCTCCCTGTG 2169
Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 2170 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCAATCCACCGAA 2229
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2230 CTCTGTGCTGGGCAATTTGGCCGAGGCACTGACAGTTGCCAGGTGCAGTGGAGTCTCT 2289
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2290 CTGTTGCTTCGAGAGGACAAATACATTTTCAAGAGTCACTTCTGGGGTCTTGGC 2349
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Db 2350 TGTGACGCCCAATAGCCCTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2409
Qy 201 GluGlyValMetArgAsnAsn 207
Db 2410 GAGGAGTGTATGAGAAATAAT 2430

RESULT 12
US-09-946-893-1
; Sequence 1, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; TITLE OF INVENTION: inhibitors
; FILE REFERENCE: Newburn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,893

```

```
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(2482)
US-09-946-893-1

Alignment Scores:
Pred. No.: 1.68e-138 Length: 2497
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-946-893-1 (1-2497)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1859 ATGCACCTTCTGGAGGACCTTGTATATCCCGAGAGTGGGTGTACCTGCTGCCCACTGC 1918

Qy 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGluValAlaHisGlnGluVal 40
Db 1919 TTGGAGAGTCCCCAAGGCCCTTCATCTACAGGTCTCTGGGTGCACACCAAGAGTG 1978

Qy 41 AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60
Db 1979 AATCTCGAACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 2038

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 2039 AAGATATTGCTTGTCTAAGCTAAGCAGTCTCTGCCGTCATCTACGACAAAGTAAATCCA 2098

Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 2099 GCTTGTCTGCCATCCCCCAATTTGTGTCTGCTGACCGGACCGAATGTTTCATCACTGGC 2158

Qy 101 TrpGluGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 2159 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 2218

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 2219 ATTGAGATAAAGTGTGCAATCGCTATGAGTTCCTGAATGGAAGAGTCCAAATCCACCGAA 2278

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2279 CTCTGTCTGGGCATTTGGCCGGAGGACCTGACAGTTGCCAGGGTGACAGTGGAGGTCT 2338

Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2339 CTGGTTTCTTCGAGAAGCAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 2398

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db 2399 TGTGACGCCCAATAAGCCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2458

Qy 201 GluGlyValMetArgAsnAsn 207
Db 2459 GAGGAGTGATGAGAAATAAT 2479

RESULT 13
US-10-135-872B-1
; Sequence 1, Application US/10135872B
; Publication No. US20040071659A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Betty
; APPLICANT: Wu, Wei Wei
```

```
; APPLICANT: McArthur, James
; APPLICANT: Patel, Salil
; APPLICANT: Jooss, Karin
; APPLICANT: Mendez, Michael
; APPLICANT: Donahue, Brian
; TITLE OF INVENTION: Viral-Mediated Delivery and In Vivo Expression of
; FILE OF INVENTION: Polynucleotides Encoding Anti-Angiogenic Proteins
; FILE REFERENCE: 3802-009-27
; CURRENT APPLICATION NUMBER: US/10/135,872B
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/287,673
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/370,634
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-135-872B-1

Alignment Scores:
Pred. No.: 1.91e-138 Length: 2732
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-135-872B-1 (1-2732)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db 1864 ATGCACCTTCTGGAGGACCTTGTATATCCCGAGAGTGGGTGTACCTGCTGCCCACTGC 1923

Qy 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGluValAlaHisGlnGluVal 40
Db 1924 TTGGAGAGTCCCCAAGGCCCTTCATCTACAGGTCTCTGCCGTCACACCAAGAGTG 1983

Qy 41 AsnLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60
Db 1984 AATCTCGAACCGCATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 2043

Qy 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 2044 AAGATATTGCTTGTCTAAGCTAAGCAGTCTCTGCCGTCATCTACGACAAAGTAAATCCA 2103

Qy 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 2104 GCTTGTCTGCCATCCCCCAATTTGTGTCTGCTGACCGGACCGAATGTTTCATCACTGGC 2163

Qy 101 TrpGluGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 2164 TGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 2223

Qy 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Db 2224 ATTGAGATAAAGTGTGCAATCGCTATGAGTTCCTGAATGGAAGAGTCCAAATCCACCGAA 2283

Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2284 CTCTGTCTGGGCATTTGGCCGGAGGACCTGACAGTTGCCAGGGTGACAGTGGAGGTCT 2343

Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Db 2344 CTGGTTTCTTCGAGAAGCAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 2403

Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTriple 200
Db 2404 TGTGACGCCCAATAAGCCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2463

Qy 201 GluGlyValMetArgAsnAsn 207
```

```
Db 2464 GAGGAGTGATGAGAAATAAT 2484
|||||
RESULT 14
US-10-193-656-1
; Sequence 1, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, TOR
; APPLICANT: HOLMDEHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / X05199
; DATABASE ENTRY DATE: 1995-03-21
; RELEVANT RESIDUES: (1)..(2732)
US-10-193-656-1
Alignment Scores:
Pred. No.: 1,91e-138 Length: 2732
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-09-992-095B-54 (1-207) x US-10-193-656-1 (1-2732)
Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
|||||
Db 1864 ATGCACCTCTGTGAGGACCTTGATATCCCAAGATGGGTGTGACTGTGCCCATGTC 1923
|||||
Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
|||||
Db 1924 TTGGAGAAGTCCCCAAGSCCTTCATCTACAAGGTCTCCTGGGTGCACCAAGAAGTG 1983
|||||
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
|||||
Db 1984 AATCTGAACCGCATGTTTCAGGAATAAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 2043
|||||
Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspLysValIlePro 80
|||||
Db 2044 AAAGATATTGCTTGTCTAAAGCTAAGCATCTCCGCTCATCTCATGCAAAAGTAATCCCA 2103
|||||
Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
|||||
Db 2104 GCTTGTCTGCCATCCCCCAAAATATATGTGTCTGCTGACCGGACCGAATGTTTTCATCACTGGC 2163
|||||
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuIysGluAlaGlnLeuProVal 120
|||||
Db 2164 TGGGGAGAAACCCCAAGGTACTTTTGGAGCTGGCCCTTCTCAAGGAAGCCCAAGCTCCCTGTG 2223
|||||
Qy 121 IleGluAsnIysValCysAsnArgTyrlsGluPheLeuAsnGlyArgValGlnSerThrGlu 140
|||||
Db 2224 ATTGAGATATAAGTGTGCATCGCTATGAGTTTCTGTAATGGAAGAGTCCATCCACCGAA 2283
|||||
Qy 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
|||||
Db 2284 CTCTGTGCTGGGCATTTGGCCGAGGCACCTGACACAGTTTCAGAGGTGACAGTGGAGGTCT 2343
|||||
Qy 161 LeuValCysPheGluIysAspLysTyrlsIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
|||||
Db 2344 CTGTTTGTCTGAGAGGACAATAATATTTCACAGAGTCACTTCTTGGGGTCTTGGC 2403
|||||
Qy 181 CysAlaArgProAsnIysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200
|||||
Db 2404 TGTGCAGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGGATT 2463
|||||
Qy 201 GluGlyValMetArgAsnAsn 207
|||||
Db 2464 GAGGAGTGATGAGAAATAAT 2484
|||||
RESULT 15
US-10-753-646-12
; Sequence 12, Application US/10753646
; Publication No. US20040138127A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS FOR INHIBITING
; FILE REFERENCE: 5940.US.P3
; CURRENT APPLICATION NUMBER: US/10/753,646
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US/08/924,287A
; PRIOR FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 08/851,350
; PRIOR FILING DATE: 1997-05-05
; PRIOR APPLICATION NUMBER: US 08/832,087
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: US 08/643,219
; PRIOR FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-753-646-12
Alignment Scores:
Pred. No.: 2,29e-138 Length: 2497
Score: 1112.00 Matches: 206
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 0
Query Match: 99.91% Indels: 0
DB: 17 Gaps: 0
US-09-992-095B-54 (1-207) x US-10-753-646-12 (1-2497)
Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
|||||
Db 1859 ATGCACCTCTGTGAGGACCTTGATATCCCAAGTGGGTGTGACTGTGCTGCCCATGTC 1918
|||||
Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
|||||
Db 1919 TTGGAGAAGTCCCCAAGSCCTTCATCTACAAGGTCTCCTGGGTGCACCAAGAAGTG 1978
|||||
Qy 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
|||||
Db 1979 AATCTGAACCGCATGTTTCAGGAATAAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 2038
|||||
Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspLysValIlePro 80
|||||
Db 2039 AAAGATATTGCTTGTCTAAAGCTAAGCATCTCCGCTCATCTACTGACAAAGTAATCCCA 2098
|||||
Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100
|||||
Db 2099 GCTTGTCTGCCATCCCCCAAAATATATGTGTCTGCTGACCGGACCGAATGTTTCTGCTCACTGGC 2158
|||||
```

```
Qy 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Db 2159 TGGGAGAGAAACCCAGGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCTGTG 2218
Qy 121 IleGluLeuLysValCysAsnArgTyrGluPheLeuLeuGlyArgValGlnSerThrGlu 140
Db 2219 ATTGAGAAATAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAA 2278
Qy 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Db 2279 CTCTGTGCTGGGCATTTGGCCGGAGGCACCTGACAGTTGCCAGGGTGACAGTGGAGGTCT 2338
Qy 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTyrGlyLeuGly 180
Db 2339 CTGGTTTCTTCGAGAGGACAAATACATTTTACAAGGAGTCACCTTCTTGGGGTCTTGSC 2398
Qy 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTyrIle 200
Db 2399 TGTGCACCCCCCAATAAGCCTGGTGTCTATGTCTGTGTTCAGGTTTGTACTTGGATT 2458
Qy 201 GluGlyValMetArgAsnAsn 207
Db 2459 GAGGAGTGTGAGNAATAAT 2479
```

Search completed: September 14, 2004, 02:26:23
Job time : 477 secs

This Page Blank (uspto)